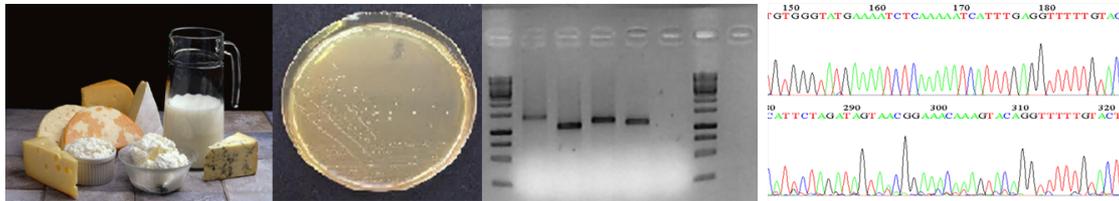




**Tesina de grado**  
**Licenciatura en Bioquímica**



**“Caracterización y análisis de la variación de *Loci*  
CRISPRs en *Streptococcus thermophilus*”**

**María Eugenia Taibo Rocha**

Tutora: Dra. María Julia Pianzola

Co-tutor: MSc. Rodrigo Achigar

Cátedra de Microbiología Molecular, Departamento de Biociencias,  
Facultad de Química, Universidad de la República

Octubre, 2016

---

## Agradecimientos

Quiero agradecer en primer lugar a María Julia por haber sido mi tutora, por dejarme formar parte de su gran grupo de trabajo, por trasmitirme sus conocimientos y por confiar en mí.

Agradecerle a mi co-tutor Rodrigo por dejarme ser parte de su proyecto, por enseñarme, por su gran paciencia y por estar siempre presente.

A todos mis compañeros y amigos del laboratorio 12, por hacerme sentir desde un principio bienvenida, por contestar mis preguntas y ayudarme siempre que lo necesito.

Agradecerles especialmente a mamá y a papá por todo su cariño, por haberme dado la mejor educación y gracias a ellos hoy soy lo que soy.

A mi hermanita Caro por todo su cariño y por hacerme sentir orgullosa de ser su hermana mayor.

A Maxi por apoyarme, por darme siempre para adelante sobre todo cuando las cosas no salían tan bien y hacer que todo sea mucho más fácil.

A la familia de Maxi por siempre estar y por dejarme ser una más de su familia.

A toda mi familia, a mis amigas y amigos de Facultad y a mis amigas de la vida.

A todos ellos,

Mil gracias!

---

# Índice

Resumen .....	1
<b>1. Introducción .....</b>	<b>2</b>
<b>1.1. Bacterias Ácido Lácticas .....</b>	<b>3</b>
Generalidades .....	3
Importancia de las bacterias ácido lácticas en la industria láctea .....	3
<b>1.2. <i>Streptococcus thermophilus</i> .....</b>	<b>7</b>
Generalidades .....	7
Metabolismo de azúcares .....	7
Sistema proteolítico .....	8
<b>1.3. Bacteriófagos .....</b>	<b>9</b>
Generalidades .....	9
Características generales de la reproducción vírica .....	9
Bacteriófagos de <i>Streptococcus thermophilus</i> .....	10
Consecuencias a nivel industrial .....	12
<b>1.4. El sistema inmune bacteriano CRISPR .....</b>	<b>13</b>
Generalidades .....	13
Proteínas asociadas y clasificación .....	15
Mecanismo de acción.....	18
Sistema CRISPR/Cas en <i>Streptococcus thermophilus</i> .....	23
La diversidad de espaciadores y su eficiencia en el mecanismo de defensa.....	25
<b>2. Objetivos y estrategia de trabajo .....</b>	<b>26</b>
<b>2.1. Objetivo general .....</b>	<b>27</b>
<b>2.2. Objetivos específicos .....</b>	<b>27</b>
<b>2.3. Estrategia de trabajo .....</b>	<b>27</b>
<b>3. Materiales y métodos .....</b>	<b>28</b>
<b>3.1. Material de partida .....</b>	<b>29</b>
Cepas bacterianas .....	29
Bacteriófagos .....	29
<b>3.2. Métodos .....</b>	<b>31</b>
Aislamiento de las cepas de <i>S. thermophilus</i> .....	31
Extracción de ADN bacteriano .....	31
Amplificación de las regiones CRISPR .....	32
Secuenciación y análisis de regiones CRISPR .....	33
Mapeo de los espaciadores en el genoma de los bacteriófagos.....	34

---

<b>4. Resultados y discusión</b> .....	35
<b>4.1 Caracterización molecular de <i>loci</i> CRISPR en <i>S. thermophilus</i></b> .....	36
<b>4.2 Comparación de los espaciadores encontrados con genomas de bacteriófagos uruguayos</b> .....	46
<b>5. Conclusiones y perspectivas</b> .....	73
<b>6. Bibliografía</b> .....	75

---

## Resumen

La fermentación es uno de los procesos de conservación de alimentos más antiguo. Su popularidad se debió a que no solo preservaba los alimentos, sino que también les otorgaba una gran variedad de sabores, formas y otras características sensoriales. Es un proceso que consiste en la transformación de materia prima simple en una gama de productos de valor añadido mediante la utilización de los fenómenos de crecimiento de microorganismos y sus actividades en diferentes sustratos.

Las Bacterias Ácido Lácticas (BALs) son el grupo de microorganismos más importante usado en la fermentación de alimentos debido a que pueden producir: sustancias antimicrobianas, polímeros de azúcar, endulzantes, compuestos aromáticos, vitaminas, enzimas y también por sus propiedades probióticas. Se destaca su capacidad para acidificar la leche mediante la producción de ácidos orgánicos, principalmente el ácido láctico. Los géneros más utilizados como cultivos iniciadores en los procesos de fermentación de productos lácteos, cárnicos y de bebidas son *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Pediococcus* y *Streptococcus*.

*Streptococcus thermophilus* es considerada como la segunda especie más importante dentro de las bacterias ácido lácticas de uso industrial. Está involucrada en el proceso de acidificación de la leche y en la elaboración de varios productos lácteos (quesos y yogurt).

A pesar de los grandes esfuerzos, la infección de bacteriófagos a cultivos de BALs iniciadoras sigue siendo la causa más común en el enlentecimiento o fermentación incompleta en la industria láctea, pudiendo causar la pérdida del total de la producción. Por lo tanto, contar con cepas resistentes a los bacteriófagos es un objetivo de interés en la industria láctea.

Se ha comprobado que las bacterias y arqueas son capaces de adquirir resistencia a bacteriófagos mediante la integración de fragmentos cortos de ácidos nucleicos foráneos en el cromosoma del huésped. Dichos fragmentos (espaciadores) son integrados en un extremo de un elemento génico repetitivo conocido como *locus* CRISPR (*clustered regularly interspaced short palindromic repeats*) [17].

Estudios de análisis genómicos comparativos en *S. thermophilus* han establecido una relación entre el *locus* CRISPR y la resistencia a bacteriófagos. Por lo tanto, contar con el conocimiento de este sistema es de suma importancia para el mejoramiento de cepas, de manera que sean más resistentes al ataque de bacteriófagos.

# 1. INTRODUCCIÓN

### 1.1 Bacterias ácido lácticas

#### Generalidades

Las bacterias ácido lácticas (BALs) son un grupo de bacterias Gram positivas, cocos o bacilos, no esporulados, anaerobias aerotolerantes y productora de ácido láctico como uno de los productos principales de la fermentación de carbohidratos [1][2]. De acuerdo a la clasificación taxonómica actual, pertenecen al phylum *Firmicutes*, clase *Bacilli*, orden *Lactobacillales*. A nivel filogenético las BALs se pueden clasificar mediante criterios de biología molecular, por ejemplo mediante la secuencia del ARNr 16S [2]. Pueden crecer a temperaturas desde 5°C a 45°C y su condición de pH óptimo es entre 5,5 y 6,5. Estas bacterias tienen un requerimiento nutricional complejo incluyendo: amino ácidos, péptidos, bases nucleotídicas, vitaminas, minerales, ácidos grasos y carbohidratos [1]. Las BALs son el grupo de microorganismos más importantes usados en la fermentación de alimentos. Tienen gran relevancia debido a características fisiológicas, tales como la utilización de sustratos, capacidades metabólicas y sus propiedades probióticas [1]. La importancia de las BALs a nivel industrial está relacionada con su aceptación como microorganismos seguros (GRAS) para el consumo humano y por sus habilidades fermentativas [1][2].

Los géneros más comúnmente utilizados como cultivos iniciadores en los procesos de fermentación de productos lácteos, cárnicos y de bebidas son *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Pediococcus* y *Streptococcus* [1]. Sin embargo, algunos géneros (*Streptococcus*, *Lactococcus*, *Enterococcus*, *Carnobacterium*) también contienen especies o cepas reconocidas como patógenos humanos o de animales, por lo que es importante contar con un conocimiento a fondo de la taxonomía, metabolismo y biología molecular de las BALs, para maximizar el aprovechamiento de las propiedades tecnológicas de estos organismos evitando riesgos potenciales [2].

#### Importancia de las bacterias ácido lácticas en la industria láctea

Las BALs están asociadas a la industria de alimentos fermentados debido a que pueden producir sustancias antimicrobianas, polímeros de carbohidratos, endulzantes, compuestos aromáticos, vitaminas, enzimas útiles y también por sus propiedades probióticas (tabla 1) [3].

La propiedad más importante es su capacidad para acidificar la leche mediante la producción de ácidos orgánicos, principalmente el ácido láctico. También se caracterizan por la producción de otros compuestos tales como; ácido acético, etanol, compuestos aromáticos, bacteriocinas, exopolisacáridos y enzimas de manera de prolongar la duración y la seguridad microbiana de los

productos fermentados (yogurt, quesos), mejorando así también la textura y contribuyendo con un perfil sensorial agradable [3].

Existen las especies homofermentativas de BAL, que convierten mayoritariamente la lactosa de la leche en ácido láctico, y por otro lado se encuentran las especies heterofermentativas que convierten la lactosa en ácido láctico, ácido acético, etanol y CO<sub>2</sub> [4].

La presencia de BALs en la fermentación de la leche puede deberse a que la leche es su hábitat natural, o por inoculación como cultivo iniciador [4]. Se entiende como cultivo iniciador una preparación microbiana de gran número de células de al menos un microorganismo que se añade a una materia prima para producir un alimento fermentado mediante la aceleración y la dirección de su proceso [3]. Aunque en las fermentaciones espontáneas el crecimiento de BALs no se puede predecir o controlar, este procedimiento se ha practicado y llevado a cabo tradicionalmente por años por ejemplo en la producción de algunos quesos artesanales.

La acción conservante de los cultivos iniciadores en los sistemas alimentarios, se atribuye a la acción combinada de una gama de metabolitos antimicrobianos producidos durante el proceso de fermentación [3], en particular productos finales, tales como algunos ácidos orgánicos (láctico, acético y propiónico). Los ácidos orgánicos pueden estar naturalmente presentes como constituyentes de los alimentos o se pueden añadir directamente o indirectamente a los productos. La fermentación de azúcar, seguido por una reducción en el pH debido a la producción de ácido láctico y otros ácidos orgánicos es un factor importante para la inhibición de crecimiento de microorganismos no deseados. El pH bajo hace que los ácidos orgánicos se vuelvan liposolubles, lo que les permite atravesar la membrana celular y alcanzan el citoplasma de patógenos. Los niveles y tipos de ácidos orgánicos producidos durante el proceso de fermentación dependen de la especie de organismos, composición del cultivo y de las condiciones de crecimiento [1].

La detección de ácidos orgánicos producidos por las bacterias del ácido láctico tiene un gran significado para la industria alimentaria ya que los cultivos iniciadores más adecuados tienen la capacidad de mejorar la calidad y la inocuidad de los alimentos. Estos se pueden seleccionar de acuerdo a sus características tecnológicas, incluida su eficacia en la producción de ácido láctico y capacidad de reducción de pH, así como la actividad antimicrobiana frente a patógenos transmitidos por los alimentos, cuando se utilizan como cultivos iniciadores [1].

Por otro lado, la aplicación de BALs como cultivos iniciadores permitió la producción y manufacturación a gran escala de productos fermentados de la leche bajo condiciones controladas.

En el caso de los quesos los cultivos iniciadores de BALs son responsables de su sabor. Según el rol de BALs estas pueden ser divididas en dos clases; las iniciadoras y las no iniciadoras. Las primeras se

## 1. Introducción

caracterizan por la producción de ácido láctico y por influir en el proceso de maduración, mientras que las segundas contribuyen más durante el proceso de maduración. La formación y caracterización del sabor de las diferentes variedades de quesos se genera durante el proceso de maduración de las dos clases de BALs, mediante los siguientes pasos: metabolismo de la lactosa, del lactato y citrato; liberación de ácidos grasos por lipólisis, degradación de caseína por proteólisis, seguido por el catabolismo de aminoácidos [5].

Tabla 1. Alimentos y bebidas fermentados, y su relación con BALs [3].

Tipo de producto fermentado	Bacterias ácido lácticas
<b>Productos lácteos</b>	
- Quesos duros sin ojos	<i>L. lactis</i> subsp. <i>lactis</i> , <i>L. lactis</i> subsp. <i>cremoris</i>
* Quesos con ojos pequeños	<i>L. lactis</i> subsp. <i>lactis</i> , <i>L. lactis</i> subsp. <i>lactis</i> var. <i>diacetylactis</i> , <i>L. lactis</i> subsp. <i>cremoris</i> , <i>Leuc. mesenteroides</i> subsp. <i>cremoris</i>
- Tipos de quesos suizos e italianos	<i>Lb. delbrueckii</i> subsp. <i>lactis</i> , <i>Lb. helveticus</i> , <i>Lb. casei</i> , <i>Lb. delbrueckii</i> subsp. <i>bulgaricus</i> , <i>S. thermophilus</i>
* Manteca	<i>L. lactis</i> subsp. <i>lactis</i> , <i>L. lactis</i> subsp. <i>lactis</i> var. <i>diacetylactis</i> , <i>L. lactis</i> subsp. <i>cremoris</i> , <i>Leuc. mesenteroides</i> subsp. <i>cremoris</i>
* Yogourt	<i>Lb. delbrueckii</i> subsp. <i>bulgaricus</i> , <i>S. thermophilus</i>
* Fermentos probióticos	<i>Lb. casei</i> , <i>Lb. acidophilus</i> , <i>Lb. rhamnosus</i> , <i>Lb. johnsonii</i> , <i>B. lactis</i> , <i>B. bifidum</i> , <i>B. breve</i>
- Kefir	<i>Lb. kefir</i> , <i>Lb. kefiranocties</i> , <i>Lb. brevis</i>
<b>Fermentos cárnicos</b>	
- Salchichas fermentadas (Europa)	<i>Lb. sakei</i> , <i>Lb. curvatus</i>
- Salchichas fermentadas (USA)	<i>P. acidilactici</i> , <i>P. pentosaceus</i>
<b>Fermentos derivados del mar</b>	<i>Lb. alimentarius</i> , <i>C. piscicola</i>
<b>Fermentos de vegetales</b>	
- Chucrut	<i>Leuc. mesenteroides</i> , <i>Lb. plantarum</i> , <i>P. acidilactici</i>
- Pickles	<i>Leuc. mesenteroides</i> , <i>P. cerevisiae</i> , <i>Lb. brevis</i> , <i>Lb. plantarum</i>
* Olivas fermentadas	<i>Leuc. mesenteroides</i> , <i>Lb. pentosus</i> , <i>Lb. plantarum</i> ,
* Vegetales fermentados	<i>P. acidilactici</i> , <i>P. pentosaceus</i> , <i>Lb. plantarum</i> , <i>Lb. fermentum</i>
<b>Salsa de soja</b>	<i>T. halophilus</i>
<b>Fermentos de cereales</b>	
* Masa fermentada	<i>Lb. sanfransiscensis</i> , <i>Lb. farciminis</i> , <i>Lb. fermentum</i> , <i>Lb. brevis</i> , <i>Lb. plantarum</i> , <i>Lb. amylovorus</i> , <i>Lb. reuteri</i> , <i>Lb. pontis</i> , <i>Lb. panis</i> , <i>Lb. alimentarius</i> , <i>W. cibaria</i>
<b>Bebidas alcohólicas</b>	
- Vino	<i>O. oeni</i>
- Vino de arroz	<i>Lb. sakei</i>

<sup>3</sup> *B.* = *Bifidobacterium*, *C.* = *Camobacterium*, *L.* = *Lactococcus*, *Lb.* = *Lactobacillus*, *Leuc.* = *Leuconostoc*, *O.* = *Oenococcus*, *P.* = *Pediococcus*, *S.* = *Streptococcus*, *T.* = *Tetragenococcus*, *W.* = *Weissella*.

Para la fabricación de yogur los cultivos de BALs más utilizados son cepas de *Streptococcus thermophilus* y *Lactobacillus bulgaricus*, los cuales mediante su asociación y beneficio mutuo generan la producción de metabolitos volátiles que determinan el sabor, liberando aminoácidos de la leche y ácidos orgánicos, produciendo más ácido láctico y compuestos aromáticos.

La leche es pobre tanto en aminoácidos libres como en péptidos cortos, y el óptimo crecimiento de las BALs requiere hidrólisis de caseína, degradación e internalización de los pequeños péptidos formados, o en su defecto la síntesis de *novo* de ciertos aminoácidos. La eficiencia del proceso de fermentación está relacionada con la capacidad para metabolizar el Nitrógeno [2].

### 1.2 *Streptococcus thermophilus*

#### Generalidades

*Streptococcus thermophilus* es considerada como la segunda especie más importante dentro de las bacterias ácido lácticas de uso industrial después del *Lactococcus lactis*. Tiene un valor de mercado de alrededor 40 billones de dólares americanos. Es una especie reconocida generalmente como segura (GRAS), alrededor de  $10^{21}$  células vivas son ingeridas anualmente por el hombre [6][7].

Posee un genoma *core* de 1271 genes [6], de los cuales algunos parecen ser originarios de otras especies lácteas tales como *Lactococcus lactis* y *Lactobacillus delbrueckii*, lo que contribuyó a su adaptación al medio lácteo[6]. Por otro lado están muy relacionados con especies de *Streptococcus* patógenos para humanos (*S.pneumoniae*, *S.pyogenes*, *S.agalactiae*), que causan por ejemplo neumonía, sepsis bacteriana o meningitis.

Esta especie está involucrada en el proceso de acidificación de la leche y en la elaboración de varios productos lácteos fermentados [8][9]. Tradicionalmente es utilizada en combinación con *Lactobacillus delbrueckii* subps. *bulgaricus* o *Lactobacillus. helveticus* para la manufacturación de yogurt y los “*hard cooked cheeses*” tales como: emmental y gruyere. Son procesos que requieren relativamente altas temperaturas (45°C). La tasa de acidificación es un rasgo dependiente del metabolismo de cada cepa, y puede estar influenciado por diversos factores como; el metabolismo lactosa-galactosa, el sistema proteolítico y de la actividad ureolítica [6].

#### Metabolismo de azúcares

*S.thermophilus* es una BALs homofermentativa obligada, con capacidad limitante de utilización de carbohidratos. Su principal función en la industria láctea es intervenir en el proceso de fermentación, convirtiendo la lactosa en lactato (L-lactato) de manera rápida y a una temperatura elevada, aunque también es capaz de contribuir en la textura y el sabor del producto final mediante la producción de otros compuestos.

A diferencia de otras bacterias Gram positivas prefiere la lactosa como principal fuente de carbono y energía en vez de glucosa. *S. thermophilus* tiene la característica de cometabolizar sacarosa y lactosa mediante un sistema de fosfotransferasas (PTS) y sustratos de no-PST respectivamente, indicando que estos carbohidratos no son jerárquicamente controlados [6]. Es incapaz de metabolizar galactosa y eso explica su presencia en el medio durante el proceso de fermentación[6].

El crecimiento de *S.thermophilus* en leche genera la producción de formiato, el cual es un compuesto importante en la estimulación del crecimiento de *Lactobacillus bulgaricus* [6].

### Sistema proteolítico

Las BALs son nutricionalmente exigentes, ya que necesitan un suministro exógeno de aminoácidos para iniciar su crecimiento [6]. El sistema proteolítico de *S. thermophilus* está compuesto por más de 20 enzimas, comprendidas por una proteasa extracelular anclada capaz de hidrolizar caseína, un sistema de transporte de aminoácidos y péptidos necesarios para la importación aminoacídica y un conjunto de peptidasas intracelulares con la función de hidrolizar péptidos de caseínas derivados de varios procesos de *house keeping* [7].

La capacidad de *S. thermophilus* para satisfacer su necesidad aminoacídica durante el crecimiento en la leche, es diferente a la de otras especies de BALs. La exigencia de aminoácido es satisfecha por la eficiencia de su capacidades biosintéticas y por la cooperación de otras especies de BALs que crecen conjuntamente en el entorno lácteo, tales como *Lactobacillus delbrueckii* subsp. *bulgaricus*. Algunas cepas de *S. thermophilus* parecen ser auxotróficas para al menos cuatro aminoácidos, Glu, Cys, His, y Met, mientras que otras BALs se sabe que son más exigentes[6]. Análisis genómicos han revelado una alta conservación en los genes relacionados al funcionamiento del proceso biosintético de aminoácidos, reflejando la importancia de la síntesis de aminoácidos para el óptimo crecimiento en la leche de esta especie[6].

### 1.3. Bacteriófagos

#### Generalidades

Los bacteriófagos son virus que infectan bacterias, y se encuentran distribuidos en todos los hábitats donde haya proliferación bacteriana. Se estima que son la entidad más ampliamente distribuida en la biósfera, con una población aproximada de  $10^{31}$  partículas virales [10]. Los bacteriófagos de ADN doble hebra y con cola, o *Caudovirales* representan el 95% de la población de bacteriófagos del planeta. Los viriones de este grupo no presentan envoltura y constan de dos partes; la cabeza y la cola. La cabeza es una carcasa proteica, presenta forma icosaédrica o derivados icosaédricos elongados, contiene una molécula de ADN doble hebra y estructuras de fijación como fibras, espículas, etc. Todas las proteínas son sintetizadas de forma independiente y luego ensambladas para formar la capsida viral[11].

Por otro lado la cola de los *Caudovirales* consiste en un tubo proteico, cuyo extremo distal se une a los receptores de la superficie celular de las bacterias susceptibles. El ADN viaja a través de ella hacia la célula bacteriana que va a infectar (etapa de inyección). El tamaño de su genoma varía entre 18 y

500 Kpb, su contenido en GC se encuentra entre el 27 y 72% y se asemeja al contenido de ADN de las células hospederas [11].

Los distintos géneros dentro del orden *Caudovirales* se diferencian por su estructura genómica (presencia o ausencia de sitios *cos* o *pac*, redundancias terminales y permutaciones circulares), presencia de genes codificantes para ADN o ARN polimerasas, ausencia o presencia de nucleótidos raros y secuencia de ADN[12].

### **Características generales de la reproducción vírica**

Los virus deben inducir a la célula hospedera a sintetizar todos los componentes necesarios para fabricar más virus. Estos componentes luego deben ser ensamblados en la estructura apropiada y los nuevos viriones deben escapar de la célula para poder seguir infectando [13].

El proceso de replicación de un bacteriófago puede ser resumido en cinco fases; la primera de unión y fijación del virus a una célula hospedadora susceptible, la segunda de penetración (inyección) del virión o de su ácido nucleico en la célula, la tercera de síntesis de ácidos nucleicos y proteínas; la cual se desarrolla desde el comienzo hasta el final de la infección, la cuarta fase es de ensamblaje de las subunidades estructurales y empaquetamiento del ácido nucleico para originar nuevas partículas víricas y por último la fase de liberación de los viriones de la célula [13].

La fase de fijación, la interacción entre el virus y el hospedero se caracteriza por ser altamente específica. Las partículas del virus contienen proteínas capaces de interaccionar con receptores (proteínas, carbohidratos, glicoproteínas, lípidos) presentes en la superficie de la célula. Los receptores determinan la susceptibilidad de la célula a ser infectada [13].

La unión de un virus a la célula origina modificaciones en el virus y/o en la célula dando lugar a la penetración, ya que para replicarse dentro de la célula es necesario el ingreso del material genético. Diferentes virus tienen diferentes estrategias de penetración, el virus envuelto es descapicado en la membrana plasmática, pero en muchos otros casos el virión entero ingresa a la célula por endocitosis [13].

Para que se dé la multiplicación viral deben surgir por replicación nuevas copias del genoma vírico y sintetizarse proteínas específicas del virus. La síntesis de estas proteínas necesita ARNm específico del virus. Al principio el virus redirige el metabolismo celular hacia la síntesis de ácidos nucleicos y proteínas víricas. Más tarde se sintetizan proteínas estructurales, que son componentes de la cubierta del virus [13].

Luego continúa la fase del ensamblaje de todos los componentes y para finalizar la liberación de los viriones maduros, el cual puede estar acompañado por la lisis celular [13].

### **Bacteriófagos de *Streptococcus thermophilus***

Todos los bacteriófagos *S. thermophilus* son miembros de la familia de *Siphoviridae* y del orden *Caudovirales*. Se caracterizan por tener una cápside isométrica de 60 nm de diámetro aproximadamente, y una cola larga no contráctil que oscila entre los 220nm y 330nm de largo [8](figura 1).



Figura 1. Microfotografía electrónica de transmisión del fago de *S. thermophilus* CGL-3 [14].

Poseen un genoma de ADN de doble hebra entre 29 y 49 kpb, y se dividen en dos grupos diferentes dependiendo del mecanismo de empaquetamiento del ADN (*cos* o *pac*) y su número de proteínas estructurales mayores. Le Marrec y colaboradores [15] encontraron una correlación estricta entre la presencia de un conjunto particular de proteínas estructurales mayores del bacteriófago y el mecanismo de empaquetamiento del ADN, demostrando que los bacteriófagos *cos* contienen dos proteínas estructurales mayores (32 y 26 kDa), y los bacteriófagos *pac* tres (41, 25, y 13 kDa) [15].

Hasta el año 2010 habían sido secuenciados un total de 11 genomas fágicos completos que comprenden ambos tipos (*cos* y *pac*). Análisis genómicos mostraron que todos los bacteriófagos poseen una organización genómica similar, identificándose diferentes regiones modulares, siendo las más conservadas los módulos de replicación del ADN y el que comprende los genes involucrados en la lisis celular [8]

Mediante análisis genómicos se individualizó un genoma *core* (entendiéndose como tal a un set de genes invariablemente presentes y conservados en un grupo de aislamientos) para los 11 bacteriófagos secuenciados y se estableció el genoma *core* de los dos tipos de bacteriófagos (*cos* y *pac*) por separado. Se concluyó que solo unos pocos ORFs están conservados en todos los bacteriófagos de *S. thermophilus*, si bien los bacteriófagos *cos* y *pac* comparten una gran cantidad de genes [8].

Los bacteriófagos de *S. thermophilus* tienen la característica de ser hospedero-específicos. Sozzi y Maret probaron 23 cepas de *S. thermophilus* contra el fago s26, y solamente una cepa fue susceptible a la infección. Otro estudio comprobó que bacteriófagos de diferentes regiones tenían distintos rangos de hospederos y que generalmente eran incapaces de replicarse en cepas de *S. thermophilus* de diferentes países [8].

### Consecuencias a nivel industrial

Estos virus bacterianos (bacteriófagos) están presentes en los ecosistemas donde se han encontrado bacterias, incluyendo nichos ecológicos generados por el hombre tales como las cubas de fermentación de alimentos. La industria ha estado tratando con este fenómeno biológico desde hace muchos años y se ha basado en una variedad de enfoques prácticos para su control, que incluyen diseños adaptados de fábrica, mejores servicios de saneamiento, ventilación adecuada, cambios en los procesos, la mejora del medio iniciador, y la rotación de cultivos. A pesar de los grandes esfuerzos, la infección de bacteriófagos a cultivos de BAL iniciadoras sigue siendo la causa más común de enlentecimiento o fermentación incompleta en la industria láctea, pudiendo causar la pérdida total de la producción [16].

En la actualidad se reconoce que la mayor fuente permanente de nuevos bacteriófagos dentro del entorno lácteo es a través de la leche cruda, con una concentración que oscila entre  $10^1$  y  $10^4$  bacteriófagos por ml [16].

Por otra parte, la concentración de bacteriófagos aumenta si se utiliza leche cruda para la fabricación de productos lácteos fermentados. En consecuencia, los bacteriófagos pueden ingresar en el proceso de fabricación y acumularse rápidamente durante la fermentación si se utilizan cepas sensibles a los bacteriófagos, alcanzando concentraciones hasta  $10^8$  unidades formadoras de placas (UFP) por ml en salmuera y hasta  $10^8$  UFP por  $m^3$  en aire. Tomadas en su conjunto, una gran diversidad de bacteriófagos está presente naturalmente en el ecosistema de la leche cruda, por tanto, la ausencia de bacteriófagos en las lecherías es inevitable [16].

Aunque la leche cruda es la fuente de bacteriófagos más lógica en el entorno industrial, varias vías de dispersión pueden ocurrir en las industrias lácteas. La aerosolización es actualmente reconocida como una importante vía de dispersión[16]. El movimiento del personal o el transporte de los equipos y/o de materias primas, puede causar la dispersión de partículas de fago en forma de aerosol. Las consecuencias de esta aerosolización son aún peores si la dispersión está restringida entre las zonas contaminadas y no contaminadas. Además, los bacteriófagos presentes en subproductos reciclados también pueden extenderse a todo el entorno de la fábrica, ya que los bioaerosoles pueden permanecer en el aire durante largos periodos de tiempo[16].

Es frecuente que la industria de productos lácteos, especialmente la fabricación de quesos, recicle concentrados de proteína del suero (CPS) para aumentar el rendimiento del producto y/o incrementar los atributos del producto final[16]. Los bacteriófagos pueden mantenerse presentes en los líquidos (suero, CPS, etc) sometidos a pasteurización y a tratamientos térmicos aún más fuertes, como por ejemplo 95 °C durante varios minutos. Por otra parte, sales, grasas, sacáridos y proteínas del suero de la leche pueden proteger a los bacteriófagos de daño térmico, lo que aumenta el riesgo de esta práctica de reciclaje. Para agravar el riesgo asociado con la CPS, el suero de la leche es frecuentemente concentrado (ultrafiltración o microparticulación), lo que aumenta los niveles de fago, debido a la posible retención de viriones por las membranas. Una recomendación general para reducir al mínimo los problemas asociados con la CPS es considerar su adición sólo para una fermentación que implique el uso de cultivos iniciadores significativamente diferentes, tales como bacterias mesófilas y termófilas. También debe tenerse en cuenta que en la fabricación de los productos derivados del suero en los cuales se usaron iniciadores naturales indefinidos (composición desconocida) no se deben añadir a los procesos impulsados por cultivos de cepas definidas (composición conocida). Los iniciadores naturales a menudo contienen bacteriófagos y los virus representan una grave amenaza para el número limitado de cepas que componen los cultivos iniciadores definidos[16].

Por otra parte, una gran diversidad de bacteriófagos de estreptococos se observa a menudo en la fabricación de queso, en contraste con una población de bacteriófagos más homogénea en las instalaciones de producción de yogur [16]. La diversidad de las poblaciones de bacteriófagos en la fabricación de queso puede ser debido a la rotación de varias cepas de *S. thermophilus* como cultivos iniciadores, en comparación con los iniciadores del yogur[16].

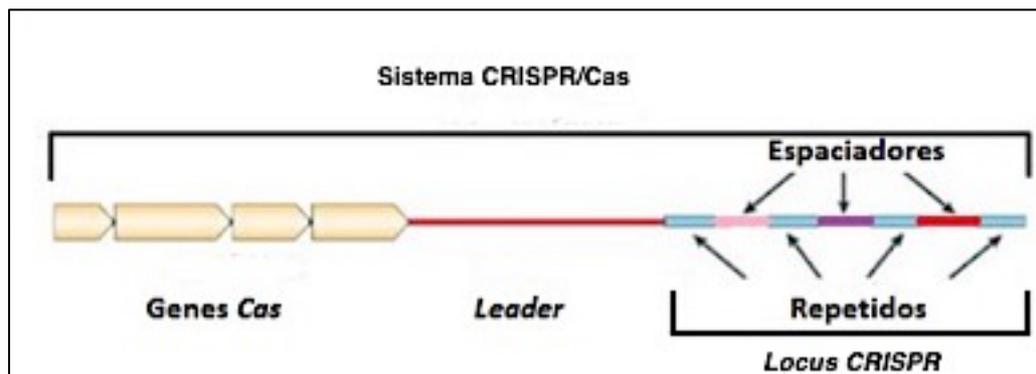
## 1.4. El sistema inmune adaptativo bacteriano CRISPR

### Generalidades

Históricamente nuestro conocimiento del sistema inmunológico microbiano recaía en los mecanismos de defensa innatos (por ejemplo, modificación de la restricción y cambios en el receptor), pero recientemente fue descubierto un sistema inmune adaptativo basado en los ácidos nucleicos [17].

Las bacterias y arqueas son capaces de adquirir resistencia a bacteriófagos y plásmidos, mediante la integración de fragmentos cortos de ácidos nucleicos extraños en el cromosoma del huésped. Dicho fragmento es integrado en un extremo de un elemento génico repetitivo conocido como CRISPR (*clustered regularly interspaced short palindromic repeat*) [18]. La inmunidad adquirida mediada por CRISPR consta de tres etapas: adquisición de ADN extraño, biogénesis CRISPR ARNs o crARNs, y por último la interferencia del objetivo. Aunque estas tres etapas básicas parecen ser comunes en todos los sistemas CRISPR, los *loci* CRISPR y las proteínas que median cada etapa de la inmunidad adaptativa son notablemente diferentes [17].

Las secuencias CRISPR son una diversa familia de repetidos de ADN que comparten una estructura en común. Cada *locus* CRISPR se compone de una serie de secuencias repetidas cortas, típicamente 20-50 pares de bases (pb) de largo, separadas por secuencias espaciadoras únicas de una longitud similar. Las secuencias repetidas dentro de un mismo *locus* CRISPR son conservadas, pero pueden diferir tanto en secuencia como en longitud en diferentes *loci* CRISPR [17] (Figura 2).



Estructura del *Locus* CRISPR

Figura 2.

Análisis filogenéticos de las secuencias de los repetidos han demostrado que los CRISPRs se pueden organizar en grupos, basándose en la similitud de las secuencias de repetidos. Algunos de estos repetidos son palindrómicos y forman ARNs con estructura en horquilla estable, mientras que otros no son capaces de formar estructuras [17].

A pesar de la extrema diversidad de secuencias de repetidos, la mayoría de los mismos tienen un motivo conservado GAAA (C/G) en el extremo 3', que puede servir como un sitio de unión para una o más de las proteínas *Cas* (CRISPR-associated) conservadas [17].

Además de la diversidad en las secuencias de repetidos y espaciadores, el número de *loci* CRISPR y la longitud de cada *locus* son variables. El número y tamaño de los *loci* CRISPR presentes en un genoma, tampoco se correlaciona con el tamaño del mismo [17].

Curiosamente CRISPR se distribuye de forma desigual entre las bacterias y arqueas. Actualmente, se han identificado *loci* CRISPR en aproximadamente el 90% de los genomas de arqueas y en el 50% de los genomas bacterianos [17].

Una secuencia rica en adenina y timina (AT) llamada líder, a menudo flanquea el *locus* CRISPR. Análisis comparativos han demostrado que las secuencias espaciadoras más cercanas al líder son más diversas, mientras que las repeticiones más alejadas (en la región conocida como *trailer*) a menudo son degeneradas [17].

Actualmente se conoce que las secuencias líder contienen elementos promotores y los sitios de unión para las proteínas reguladoras fundamentales para la expresión de crARNs y para la adquisición de nuevas secuencias [17].

### Proteínas asociadas y clasificación

Análisis comparativos han identificado un *cassette* variable de los genes *cas*, que normalmente se encuentra adyacente a un *locus* CRISPR. Cuatro genes *cas* se identificaron inicialmente en los genomas que contenían CRISPRs, pero la acumulación de secuencias de genomas y la implementación de métodos de búsqueda cada vez más sofisticados han permitido identificar aproximadamente 45 familias de genes diferentes que se encuentran comúnmente en asociación con CRISPRs [17]. Seis de estos genes *cas* (*cas1-cas6*) están ampliamente conservados y son considerados genes *cas core*. Solo *cas1* y *cas2* están universalmente conservados en los genomas que contienen *loci* CRISPR. *cas1* es característico de este sistema inmune, y análisis filogenéticos de las secuencias de *cas1* sugieren que existen versiones distintas del sistema CRISPR [17]. Cada uno de estos filotipos se define con una composición única y un ordenamiento característico de los genes *cas* (Figura 3). Esta clasificación basada en estos genes parece correlacionarse muy bien con otra clasificación basada en la secuencia de los repetidos, lo que sugiere que las proteínas *Cas* interactúan específicamente con un determinado *loci* CRISPR [17].

# 1. Introducción

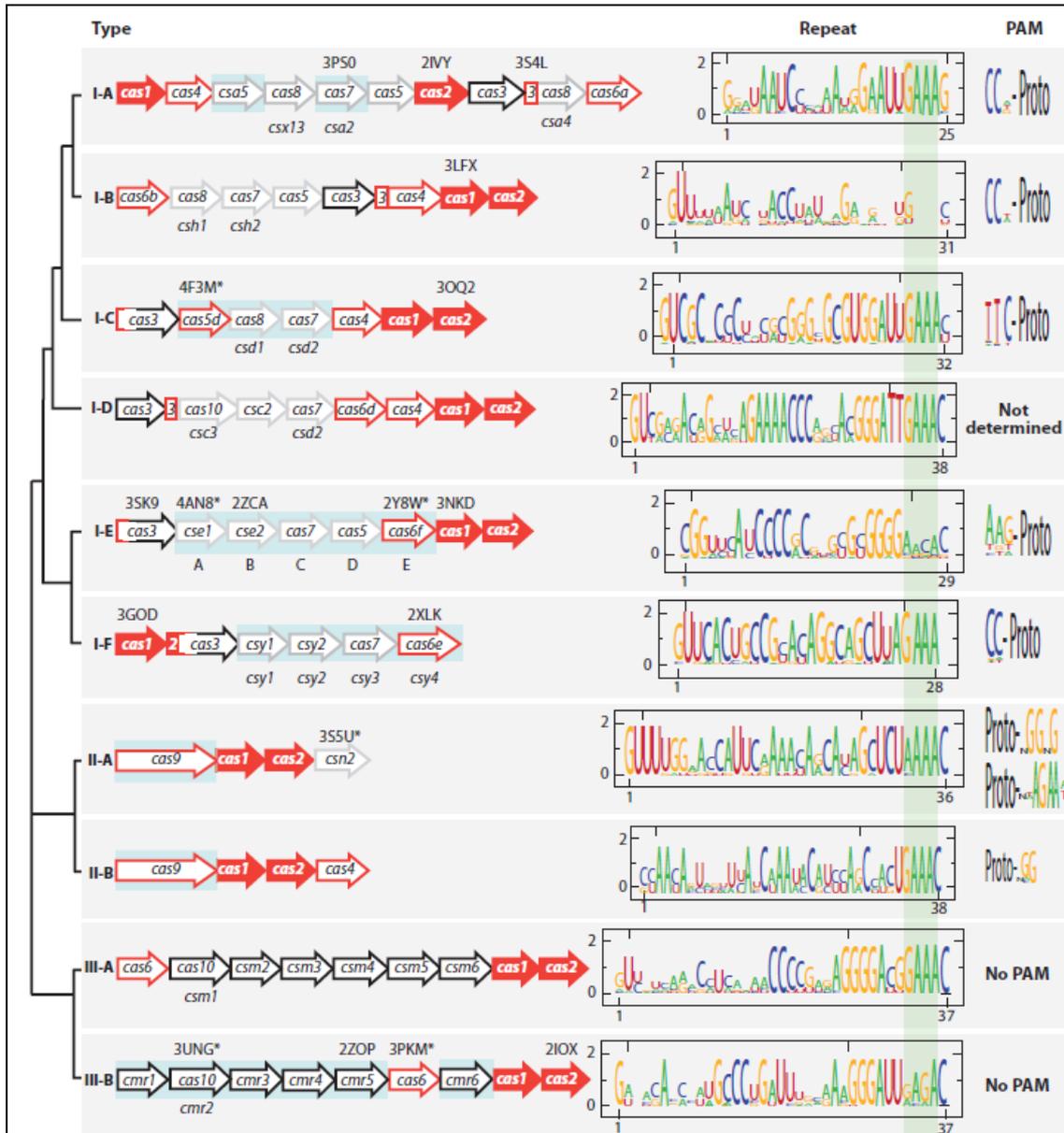


Figura 3. Diferentes tipos y subtipos de sistemas CRISPR. Se observa en la figura las diferentes organizaciones de los genes *cas* y las secuencias espaciadoras [17].

Se han determinado tres tipos de sistemas CRISPR/Cas principales, y cada uno de estos tipos a su vez pueden dividirse en subtipos (es decir, tipo IA-F, tipo IIA-B, y tipo III A-B) [17].

Nuevos análisis proporcionados por Makarova y colaboradores, 2015, sugieren que los sistemas CRISPR-*cas* pueden ser divididos sobre la base de los genes que codifican los módulos efectores. De este modo, se introduce un nuevo y más amplio nivel de clasificación de los sistemas de CRISPR-*cas*, dividiéndolos en dos clases. Los sistemas de clase 1 poseen múltiples subunidades de complejos crARN-efectores, mientras que en los sistemas de clase 2, todas las funciones del complejo efector son llevadas a cabo por una sola proteína, como por ejemplo Cas9. También encontraron evidencia de dos nuevos tipos putativos, tipo IV y tipo V, que pertenecen a la clase 1 y clase 2,

respectivamente. Estas observaciones dan lugar a un nuevo sistema de clasificación en la que los sistemas de CRISPR-*cas* se agrupan en cinco tipos, cada uno con una composición distintiva de módulos de expresión, de interferencia y de adaptación (Figura 4). Estos cinco tipos se dividen en 16 subtipos, incluyendo cinco nuevos subtipos (II-C, III-C y III-D, junto con los subtipos individuales de tipo IV y los sistemas de tipo V)[19].

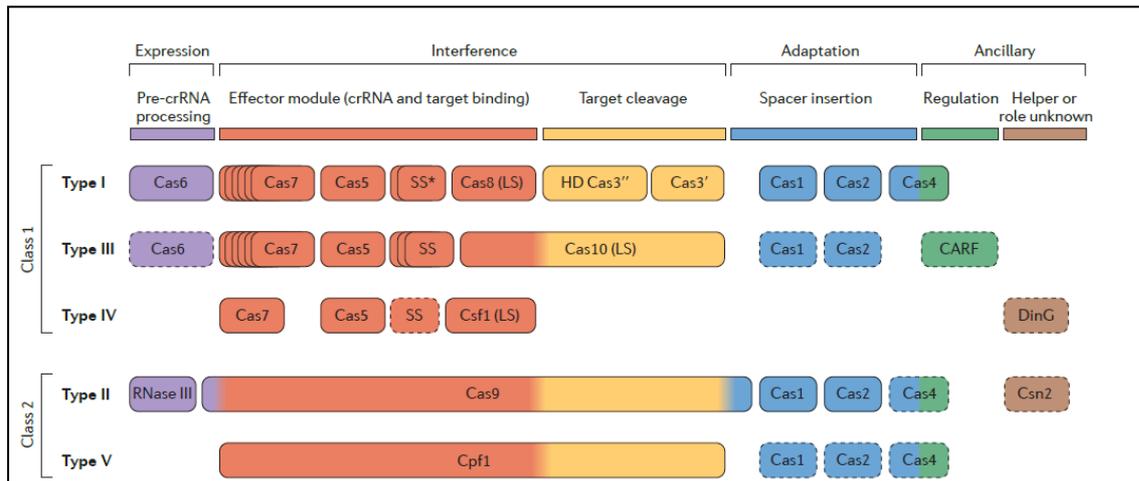


Figura 4. Clasificación funcional de las proteínas Cas [19].

Los cuatro tipos I, II, III y V poseen dos genes universales: *cas1*, el cual codifica una DNasa sin especificidad de secuencia, la cual podría estar involucrada en la incorporación de nuevos espaciadores al sistema [20] y *cas2*, el cual codifica una endorribonucleasa metal-dependiente [21]. Sin embargo, difieren substancialmente en el resto de los genes que los componen, y cada uno se caracteriza por poseer un gen marcador único. Los genes marcadores para los cuatro tipos son respectivamente: *cas3* (una helicasa de la superfamilia 2, con un dominio nucleasa N-terminal) *cas9* (una proteína de gran tamaño con un dominio *RuvC-like* y un dominio HNH nucleasa), *cas10* (una proteína que contiene un dominio homólogo al dominio “palma” de las ADN polimerasas) [19] y *Cpf1* (una proteína grande que contiene un dominio *RuvC-like* nucleasa homólogo al dominio de *Cas9*, pero carece del dominio HNH nucleasa)[19].

Por último en esta clasificación se encuentra el tipo IV que carece de los genes *cas1* y *cas2* y se caracteriza por poseer el gen *Cfs1* que codifica la subunidad mayor y pertenece a la familia de las proteínas *Cas8* [19].

## Mecanismo de acción

El mecanismo de protección de cada uno de los sistemas CRISPR/*cas* se puede dividir en tres etapas generales; adaptación (adquisición de nuevos espaciadores), Expresión de *cr*ARN e interferencia (Figura 5).

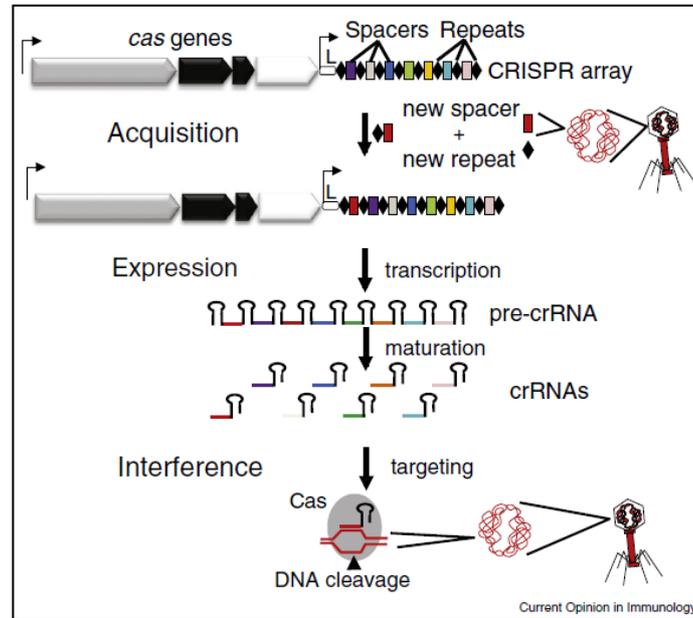


Figura 5. Sistema inmune CRISPR/*cas* [22].

## Adaptación

En el 2005, estudios independientes informaron que las secuencias espaciadoras dentro de los *loci* CRISPR eran a menudo idénticas a las secuencias de bacteriófagos y plásmidos [17].

Barrangou y colaboradores, probaron la hipótesis de que CRISPR es parte de un sistema inmune adaptativo, desafiando a una cepa industrial de *S. thermophilus* con dos bacteriófagos diferentes aislados de muestras de yogur y luego investigaron estos cultivos de bacteriófagos mutantes resistentes [17]. De acuerdo con observaciones bioinformáticas anteriores, se añadieron todas las nuevas secuencias de espaciadores de manera polarizada en el extremo líder del *locus*, y la adición de una secuencia de repetidos acompañaba la integración de cada espaciador. Naturalmente las secuencias espaciadoras adquiridas se derivan de cualquiera de las hebras del ADN invasor, y el número de nuevos espaciadores se correlaciona con el grado de resistencia al fago. Es importante destacar que, el polimorfismo de un solo nucleótido entre el espaciador y la secuencia del fago no proporciona la resistencia, lo que sugiere que la secuencia del espaciador es crítica para la protección [17].

La fase de adaptación ofrece la memoria genética que es un requisito previo para la expresión y las subsiguientes fases de interferencia que neutralizan los ácidos nucleicos que vuelven a invadir. La inserción de nuevos espaciadores se ha demostrado experimentalmente en varios subtipos CRISPR-Cas; el Tipo IA (*Sulfolobus solfataricus* y *Sulfolobus islandicus*), IB (*Haloarcula hispanica*), IE (*E. coli*) y IF (*Pseudomonas aeruginosa* y *Pectobacterium atrosepticum*) y el Tipo II-A (*S. thermophilus* y un sistema de *Streptococcus pyogenes* expresado en *Staphylococcus aureus*). Hay dos tipos de adquisición de espaciadores; *naïve*, cuando el invasor no se ha encontrado previamente, y *primed*, cuando hay un registro pre-existente del invasor en el *locus* CRISPR [23]

Aunque la adquisición de espaciadores se ha observado, su mecanismo todavía no se conoce totalmente. Conceptualmente el proceso se puede dividir en dos pasos: la selección del protoespaciador y la generación del espaciador seguido de su integración en el *locus* CRISPR y la síntesis de un nuevo repetido [23].

Ocasionalmente la eliminación de espaciadores es requerida para limitar el tamaño del *locus*. Los factores claves en la incorporación de espaciadores son las proteínas *Cas1* y *Cas2* [21]. Esto fue comprobado sobreexpresando ambas proteínas del Tipo I-E en *E. coli*, lo cual resultó en la integración del espaciador, incluso en la ausencia de las otras proteínas *Cas* [21]. Ambas *Cas1* y *Cas2* son nucleasas y mutaciones en el sitio activo de *Cas1* impiden la integración del espaciador en *E. coli* [23]. *Cas1* y *Cas2* de *E. coli* forman un complejo donde un dímero de *Cas2* interacciona con dos dímeros de *Cas1*. La formación del complejo es requerido para la adquisición del nuevo espaciador, pero la actividad nucleasa de *Cas2* es necesaria. *Cas1* preferentemente se une al ADN del *locus* CRISPR de una manera dependiente de *Cas2*, cumpliendo un rol directo en la adquisición del espaciador [21]. Se puede especular que el complejo *Cas1-Cas2* transporta el espaciador y permite su integración. Otros factores que se sabe que son requeridos para la adición del espaciador son: *Cas9*, *Csn2* y *tracrRNA* en el Tipo II-A y *Cas4* en el Tipo I-B [23]. El rol de *Csn2*, *Cas4* y *tracrRNA* no está claro hasta el momento, pero probablemente *Cas9* guía la maquinaria necesaria para la integración. Polimerasas, ligasas y proteínas de recombinación es probable que realicen los pasos genéricos en el proceso de adaptación, ya que dichos factores pueden ser encontrados en cada célula huésped [23].

La selección del espaciador parece guiarse por ciertos elementos presentes en la secuencia diana. Análisis de secuencias han revelado un motivo corto al lado de las secuencias proto-espaciadoras (secuencias de ADN extraño seleccionadas para la integración) llamada *protospacer adjacent motif* (PAM), que es crucial para la discriminación entre lo propio y no propio [17].

La importancia de los motivos PAMs va más allá de la selección del proto-espaciador. Los plásmidos o bacteriófagos que tienen una sola mutación en el PAM ya no son sensibles a la interferencia mediada por CRISPR, incluso cuando las secuencias espaciadoras y proto-espaciadora son totalmente complementarias [17].

En ambos sistemas Tipo I-E y Tipo II-A, se demostró que una parte de la región *leader* y un repetido son requeridos para la integración del espaciador. El repetido más próximo a la región *leader* sirve como molde para la síntesis de nuevos repetidos [23].

Curiosamente, la integración de espaciadores *in vitro* también se puede realizar en otras secuencias predichas para formar estructuras cruciformes, en ausencia de secuencias repetidas. Tomados en conjunto, la integración del espaciador está dirigida tanto por la secuencia y como de la estructura del *locus* CRISPR [23].

### Expresión

La transcripción del *loci* CRISPR-*Cas* para generar el complejo ARN-proteína guía sigue un mecanismo general en la mayoría de los microorganismos, aunque presenta diferencias tipo-específicas. Todos los sistemas transcriben el *locus* CRISPR; procesan el ARN con proteínas *Cas* ribonucleasas y forman el complejo CRISPR-riboproteína (crRNP) [23].

La secuencia líder contiene elementos promotores y sitios de unión para proteínas reguladoras, además de los elementos importantes para la adquisición de espaciadores [23].

Un transcripto primario de gran tamaño, el pre-crARN, es generado y puede contener una serie de estructuras secundarias (heparina) si el *locus* CRISPR contiene repetidos palindrómicos. El pre-crARN es procesado en pequeñas unidades correspondientes a un solo espaciador seguido de una secuencia parcial de un repetido. La proteína *Cas* responsable del procesamiento y si forma parte del complejo guía varía en cada subtipo (Figura 5). Incluso cuando los tres tipos de CRISPR-*Cas* coexisten no intervienen en el procesamiento del pre-crARN de los otros [23]. La comparación general de los tres tipos de sistemas CRISPR-*Cas* presentan similitudes en el procesamiento del pre-crARN como en la estructura formada por el complejo crARN. Todos los sistemas Tipo I y Tipo III utilizan la proteína *Cas6* para el procesamiento del pre-crARN, excepto el Tipo I-E que utiliza la proteína *Cas5d*. Tanto *Cas6* como *Cas5d* pertenecen a una familia diversa de endorribonucleasas-CRISPR específicas que exclusivamente interactúan con las secuencias de repetidos de sus *loci* CRISPR asociados [17]. Por ejemplo, para el sistema Tipo III la maduración del pre-crARN también involucra un paso en el procesamiento secuencia-específico mediado por *Cas6*, seguido por una secuencia no específica-

crRNA basado en la regla de corte en el extremo 3' para producir crRNAs maduros con un extremo 5' definido y un extremo 3' variable (figura 5) [23].

El sistema Tipo II por el contrario emplea un mecanismo diferente para la biogénesis del crARN, en el cual el procesamiento es dependiente de la RNasa III codificada por el hospedero y de un tracrARN (ARN trans pequeño) que interactúa mediante complementariedad de bases con el pre-crARN. Además el Tipo II requiere de la proteína *Cas9* para el procesamiento (figura 6) [23].

El sistema tipo II es el sistema activo en *S. thermophilus* y se explicará en detalle más adelante.

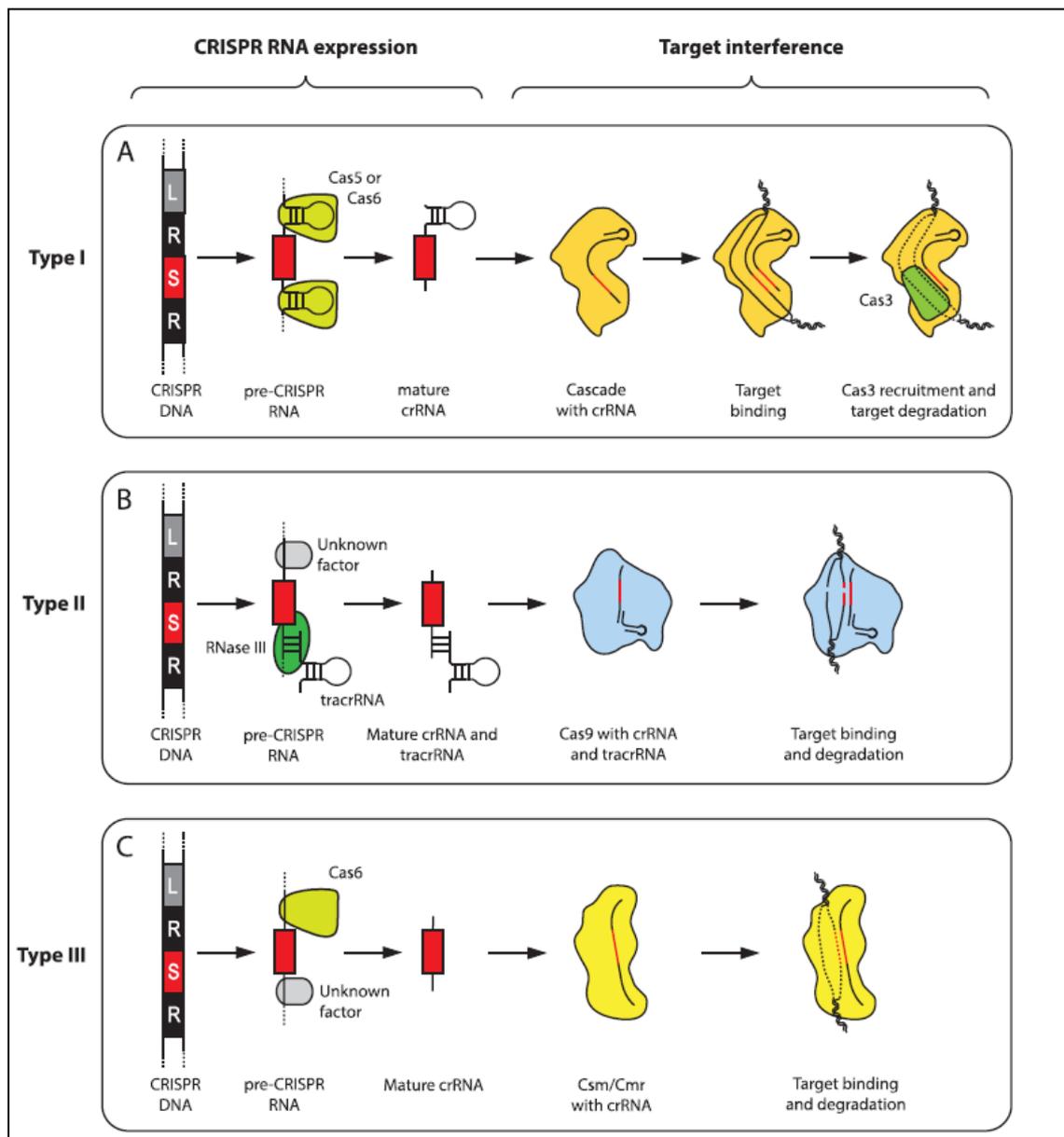


Figura 6. Modelo del procesamiento del crARN e interferencia de los diferentes Tipos de sistemas CRISPR-Cas [23]

### Interferencia

El reconocimiento del objetivo para todos los sistemas de CRISPR implica la hibridación de la secuencia de crRNA-espaciadora con una secuencia de ADN complementaria del invasor. Esto presenta un problema potencial para los sistemas de CRISPR que se dirigen a ADN de doble cadena (tipo I y II) o para los que atacan ARN (tipo III-B), debido a que las estructuras secundarias del ARN pueden obstruir sitios de unión al objetivo [17].

El complejo de reconocimiento guiado por crRNA de *E. coli*, llamado *Cascade*, se une preferentemente a dsDNA largos (plásmido o ADN de fago) que está superenrollado negativamente. El superenrollamiento negativo compacta dsDNA, por lo tanto, secuencias separadas por largas distancias a lo largo del DNA pueden ser colocadas en estrecha proximidad, acelerando el proceso de búsqueda para algunas proteínas vinculadas al dsDNA. Además, el superenrollamiento negativo introduce tensión torsional que facilita la separación de las cadenas [17].

Los primeros 8 nucleótidos del espaciador crRNA son los más importantes para la unión. Esta región del espaciador crRNA se conoce como la secuencia semilla, y mutaciones de un solo nucleótido en esta región evitan que se produzca la unión. En contraste, múltiples mutaciones en regiones diferentes a esta no impiden que la unión se realice con alta afinidad y se mantiene su eficiencia durante el desafío del fago [17].

El motivo PAM anteriormente descrito es necesario tanto para la adquisición de espaciadores como para la correcta interferencia. Ya que el mismo es identificado por el complejo de reconocimiento evitando la autodestrucción del ADN del hospedador [17].

Las proteínas involucradas en el proceso de degradación del ADN extraño es diferente para cada tipo de sistema (figura 5) [17]. Por ejemplo en el sistema tipo II únicamente es necesaria la proteína *Cas9* para llevar a cabo el procedimiento [17].

### Sistema CRISPR/Cas en *Streptococcus thermophilus*

Existen hasta ahora cuatro sistemas CRISPR/Cas presentes en *S. thermophilus*. CRISPR 1 y CRISPR 3 son del tipo IIA, CRISPR 2 es del tipo IE y CRISPR 4 es del tipo IIIA. Con respecto a su actividad, con esto se hace referencia a la habilidad de integrar nuevos espaciadores en respuesta a bacteriófagos, CRISPR 1 es el que muestra mayor actividad, seguido por CRISPR 3 [9].

Detallaremos el sistema de acción para el tipo IIA, ya que es el que presentan los *locus* CRISPR con mayor actividad (figura 7).

## 1. Introducción

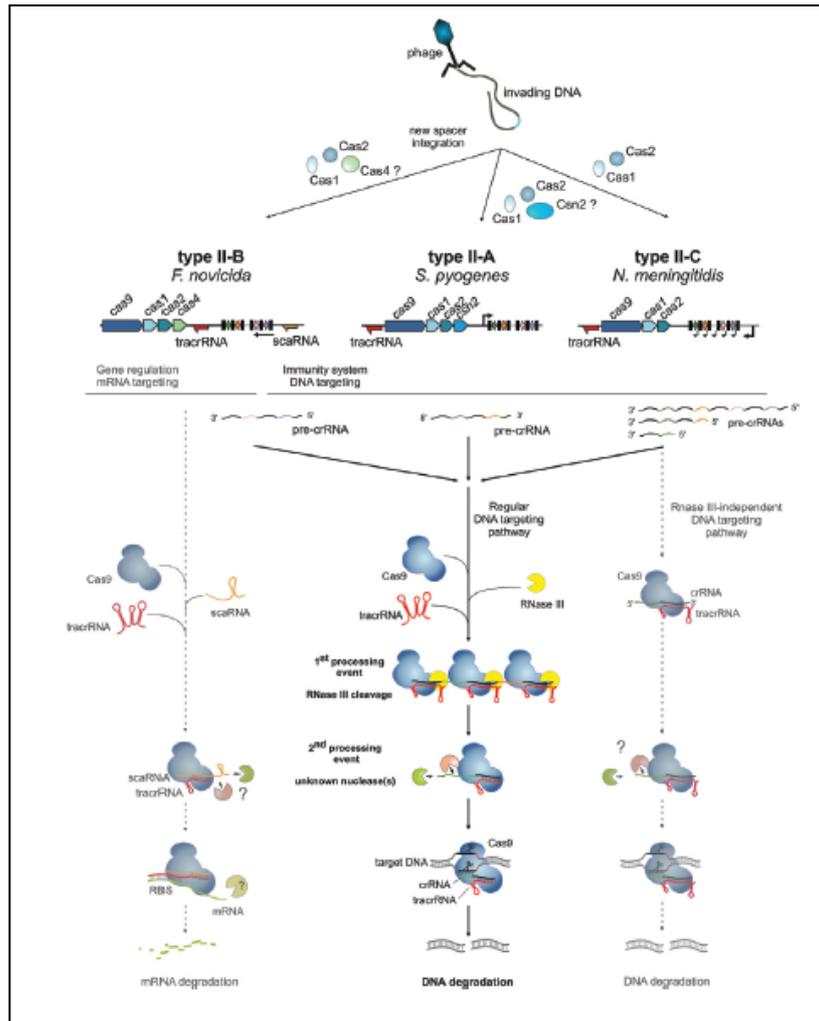


Figura 7. Esquema general del mecanismo de los sistemas CRISPR-cas tipo II [26].

Los sistemas tipo II constan de cuatro genes básicos: *cas9*, *cas1*, *cas2* y *cas4* (subtipo IIA) o *cas9* (subtipo IIB). El gen *cas9*, codifica una gran proteína multimérica, que participa tanto en la biogénesis del crARN como en la destrucción del ADN extraño (dos dominios HNH y RuvC-like nucleasas) [26].

La biogénesis de los crARNs es muy particular ya que requiere de un segundo ARN trans activante (tracrARN), el cual se encuentra aguas arriba y en la hebra opuesta al locus CRISPR. Todos los ortólogos tracr-ARN se caracterizan por la presencia de una secuencia anti-repetido homóloga a las secuencias de repetidos del locus CRISPR [26].

El primer evento de procesamiento implica el apareamiento de bases del tracrRNA con la secuencia repetida del pre-crRNA en presencia de Cas9, formando un dúplex de ARN el cual es reconocido y cortado por la RNasa III bacteriana. El proceso de maduración continúa mediante la acción desconocida aún de determinadas nucleasas. La proteína Cas9 es la encargada de guiar y degradar mediante cortes el ADN invasor [26].

### La diversidad de espaciadores y su eficiencia en el mecanismo de defensa

De acuerdo a Lopatina et al., 2016 existe diversidad natural de los espaciadores dentro de las poblaciones bacterianas y ésta es muy alta. El beneficio de la diversidad de espaciadores se explica en un estudio reciente de van Houte 2016 en el cual, utilizando ensayos de infección demostró que los fagos no pueden superar el sistema CRISPR-Cas de defensa mediante mutaciones puntuales en sus genomas cuando la diversidad de espaciadores en la población bacteriana es suficientemente alta [27]. Por tanto, es de gran interés estudiar el contenido de los espaciadores de los *loci* CRISPR en las cepas de uso comercial, su funcionamiento y su ubicación en el genoma de los bacteriófagos de forma de desarrollar nuevas estrategias que conduzcan al mejoramiento en su resistencia a fagos de cepas de *S. thermophilus* a nivel industrial.

Según el criterio del Semenova [29] para que el espaciador funcione adecuadamente y le otorgue resistencia a la cepa, debe cumplir con las siguientes condiciones; la secuencia protoespaciadora (se entiende como secuencia protoespaciadora a la secuencia del espaciador ubicada en el genoma del fago) debe tener un motivo PAM adyacente en el extremo 3' característico para cada sistema, dicho motivo no debe presentar mutaciones y por último, los 8 nucleótidos (secuencia semilla) más próximos al motivo PAM del protoespaciador deben complementarse perfectamente con la secuencia del espaciador del *locus* CRISPR de la cepa.

## **2. OBJETIVOS Y ESTRATEGIA DE TRABAJO**

### 2.1. Objetivo general

Caracterizar la variación de *loci* CRISPRs en las cepas de *S. thermophilus*, con el fin de aportar la información necesaria para la construcción de cepas resistentes a bacteriófagos de utilidad tanto para la industria láctea nacional como mundial.

### 2.2. Objetivos específicos

- Caracterización molecular de cepas *S. thermophilus* mediante el análisis del contenido de los *loci* CRISPR
- Comparación de los espaciadores encontrados en los *loci* CRISPR con genomas de bacteriófagos uruguayos

### 2.3. Estrategia de trabajo

Se amplificarán por reacciones de PCR, secuenciarán y analizarán los diferentes *loci* CRISPR de las cepas de la colección de *Streptococcus thermophilus* de Facultad de Química. El contenido nucleotídico de los espaciadores de las cepas estudiadas será evaluado mediante comparación entre cepas y análisis de bases de datos, de forma de determinar cuál es el potencial de resistencia que confiere cada espaciador a bacteriófagos, a plásmidos, etc. A su vez se mapearán los distintos espaciadores encontrados en las cepas estudiadas en genomas de bacteriófagos uruguayos, previamente obtenidos por nuestro grupo de trabajo, los cuales no están presentes en las Bases de datos de GenBank.

### **3. MATERIALES Y MÉTODOS**

#### 3.1. Material de partida

##### Cepas bacterianas

Un total de 50 cepas de *Streptococcus thermophilus* fueron utilizadas en este trabajo (tabla 2). La colección de cepas fue suministrada por la empresa distribuidora de fermentos lácticos BIOMÁS S.A. La colección se conserva en el freezer de -80 de la cátedra de Microbiología de Facultad de Química.

##### Bacteriófagos

El material de partida fueron 15 bacteriófagos aislados en un trabajo previo [28] a partir de muestras provenientes de plantas industriales y potes de fermento. Estos provienen del análisis de un total de 432 muestras (la mayoría suero) colectadas de plantas de producción de queso nacionales entre los años 2010 y 2012, de las cuales 10 corresponden a potes de fermento provenientes de Argentina proporcionados por la empresa BIOMÁS S.A.

### 3. Materiales y métodos

Tabla 2. Cepas utilizadas en este estudio y matriz con resultados de ensayos de sensibilidad/resistencia de las cepas con los 5 bacteriófagos diferentes aislados en Uruguay [28]

Cepas	Bacteriófagos				
	p53	p73	p93	p107	p128
TH4	+	-	-	-	-
En 7	-	-	-	-	+
Diag 2	-	+	-	-	-
CHE 102	-	-	-	+	-
Bio 4	-	-	+	-	-
En6	-	-	-	+	-
St 307	+	-	-	-	-
St sac 2	-	-	-	-	-
En8	-	-	-	+	-
CVA 53	-	-	-	-	-
Fcia 3	-	-	-	-	-
Diag 16	-	-	-	-	-
LF 505	-	-	-	-	-
LF 3	-	-	-	-	-
T	+	-	-	-	-
Diag 12	-	-	-	-	-
LF 507	-	-	-	-	-
St 2	-	-	-	-	-
.19	+	-	-	-	-
11 CH	-	-	-	-	-
G 72	-	-	-	-	-
Cr 1	+	-	-	+	-
Diag 9	+	-	-	-	-
St 62	+	-	-	-	-
pot	+	-	-	-	-
.3	-	-	-	+	-
En 24	-	-	-	+	-
St 179	-	-	-	+	-
St 43	-	-	-	+	-
LF 1	-	-	-	-	-
LF 2	-	-	-	-	-
Fcia 2	-	-	-	-	-
St 309	-	-	-	-	-
Hn	-	-	-	-	-
En 1	-	-	-	-	-
CVA 52	-	-	-	-	-
LF 502	-	-	-	-	-
TA	-	-	-	-	-
St 176	-	-	-	-	-
En 5	+	+	-	+	-
Diag 5	+	-	-	-	-
F 81	-	+	-	-	-
11L	-	+	-	-	-
EL	+	-	-	-	-
St 172	+	-	-	-	-
St 175	-	-	-	-	-
En 10	+	-	-	-	-
.36	-	-	-	-	+
Diag 1	-	-	-	-	-
Tma	-	+	-	-	-

## 3.2. Métodos

### Aislamiento de las cepas de *S. thermophilus*

Previo al inicio del trabajo fue necesario verificar que las cepas madres se encontraran puras. Se tomó del freezer de -80 una alícuota de 10µl de cada cepa madre, se realizó un aislamiento en medio M17 agar por 48 horas a 37°C, luego se subcultivó una colonia característica en medio M17 agar por 24 horas a 37°C, por último se realizó una tinción simple y se observó al microscopio de manera de corroborar que presentaba la morfología adecuada, cocos en cadena, para descartar posibles contaminantes. Si la cepa se encontraba pura se proseguía con la extracción de ADN, amplificación de los *loci* CRISPR, secuenciado y análisis. De lo contrario se repetía el procedimiento anterior.

### Extracción de ADN bacteriano

El objetivo de este protocolo es obtener ADN total a partir de cultivos puros de *S. thermophilus*. Este es un método estándar de extracción de ADN para microorganismos Gram positivos. Se utiliza bromuro de hexadeciltrimetilamonio (CTAB), un detergente catiónico con la propiedad de precipitar ácidos nucleicos en soluciones de baja concentración iónica, permaneciendo las proteínas y polisacáridos neutros en solución. En soluciones con altas concentraciones iónicas, el CTAB precipitará proteínas y polisacáridos neutros, dejando los ácidos nucleicos en solución. Se centrifugan 5ml de cultivo puro (ON) de *S. thermophilus* por 5 minutos 6000 rpm. Se descarta el sobrenadante y se resuspende el pellet en 0,5ml de buffer TE25S. Se añaden 10µl de solución de lisozima, se mezcla suavemente e incuba a 37° durante 30 minutos. Luego se añaden 5µl de solución de proteinasa K y 30µl de SDS 10%, se mezcla por inversión e incuba a 55° durante 60 minutos. Es necesario mezclar por inversión ocasionalmente durante este tiempo. Se añaden 100µl de NaCl 5M y nuevamente se mezcla por inversión, se añaden 65µl de solución CTAB/NaCl y se mezcla vigorosamente e incuba a 55°C durante 10 minutos. La muestras se dejan termostatar a 37°, inmediatamente se le añaden 500µl de cloroformo:alcohol isoamílico (24:1) y se agita orbitalmente durante 30 minutos. Se centrifuga a máxima velocidad (13000 rpm) durante 15 minutos, se transfiere el sobrenadante a un tubo limpio, se le añaden 0,6 volúmenes de isopropanol y se mezcla por inversión. Seguidamente se centrifuga a máxima velocidad a 4°C durante 10 minutos, se lava el pellet con etanol 70%, se seca y resuspende en 50µl de agua MQ. Para terminar se añaden 1µl de RNasa (1mg/ml), se incuba durante 15 minutos a 37°, y por último se guardar a -20°C

El ADN obtenido se visualiza mediante electroforesis en gel de agarosa al 0,8% con bromuro de etidio (0,5µg/ml). El resultado esperado es una única banda en la parte superior del gel (por encima de la banda más grande del marcador de peso molecular Smart Ladder Eurogentec). El grosor de la banda permite cuantificar la cantidad de ADN obtenido mediante comparación con las bandas del marcador en las condiciones establecidas por el fabricante. En caso de observarse *smears* o no verse banda, el protocolo debe repetirse.

#### **Amplificación de las regiones CRISPR**

Las amplificaciones de las regiones CRISPR de las cepas en estudio de *S. thermophilus* se realizaron usando dos sets de *primers* específicos (tabla 3). Cada set de *primers* está dirigido a una de las dos regiones CRISPR diferentes detalladas en la tabla. Este no es un ensayo tipo multiplex, cada región se amplifica preparando una mix diferente. Se detalla el protocolo utilizando la enzima *Taq DNA Polymerase* de la empresa Feldan y *Phusion High-Fidelity DNA Polymerase* de la empresa Thermo Scientific. Los primers fueron diseñados por el equipo del Dr. Sylvain Moineau en la Université Laval, Quebec, Canadá [26].

Para las reacciones de PCR convencionales el volumen final de la reacción fue de 50µl. La reacción de amplificación del *locus* CRISPR 1 contiene: 5 µl 10X PCR buffer, 3,2µl dNTPs 10mM, 2µl de cada primer 10µM (CR1-up y CR1-down), 1µl MgCl<sub>2</sub>, 2µl ADN y 33,4 µl agua MQ. Y la reacción de amplificación del *locus* CRISPR 3: 10µl 5X PCR buffer, 2µl dNTPs 5mM, 1,5µl DMSO, 2µl de cada primer 5mM (CR3-up y CR3-Trail2R), 0,5µl Phusion, 2µl de ADN y 30 µl agua MQ.

El programa de temperaturas para la primera reacción fue el siguiente: un primer paso de desnaturalización a 94°C por 5 minutos, seguido de 35 ciclos de desnaturalización (94°C por 30s), hibridación (58°C por 30s) y extensión (72°C por 1:30 minutos). Luego se realizó una etapa de extensión a 72°C por 10 minutos y por último una etapa a 4°C por 10 minutos. Para la segunda reacción el programa fue el siguiente: una etapa de desnaturalización a 98°C por 30 segundos, seguida de 35 ciclos de desnaturalización (98°C por 10s), hibridación (59,8°C por 17s) y extensión (72°C por 45s). Luego se realizó una etapa de extensión a 72°C por 10 minutos y por último una etapa a 4°C por 10 minutos.

Con el mismo fin se realizaron reacciones por real time PCR, el volumen final de la reacción fue de 24µl, conteniendo: 1µl de cada primer 5mM, 12,5µl 10X SYBR Green RotorGene master, 2µl de ADN, 8,5µl agua MQ. La reacción se llevó a cabo en un equipo Rotor-Gene 6000 (Corbett Research). Las condiciones establecidas fueron las siguientes: primero una etapa de desnaturalización a 94°C por 15s, seguida por 40 ciclos de desnaturalización (94°C por 30s),

hibridación (54°C por 30s) y extensión (72°C por 125s). Las condiciones de amplificación para ambos *loci* CRISPRs son las mismas

Los productos de PCR se visualizaron mediante electroforesis en gel de agarosa al 0,8% con bromuro de etidio (0,5µg/ml). El tamaño de las regiones fue variable, por lo que en algunos casos se debieron ajustar los tiempos de extensión para obtener resultados óptimos.

Tabla 3. Cebadores utilizados para la amplificación de los *loci* CRISPR 1 y CRISPR3

Cebador	Secuencia	Tm	Referencias
CR1-up	5'-TGCTGAGACAACCTAGTCTCTC- 3'	64°C	[28]
CR1-down	5'-TAAACAGAGCCTCCCTATCC- 3'	61°C	[28]
CR3-up	5'-GGTGACAGTCACATCTTGCTAAAACG- 3'	66°C	[28]
CR3-Trail2R	5'-GCTGGATATTCGTATAACATGTC- 3'	60°C	[28]

#### Secuenciación y análisis de regiones CRISPR

El objetivo de estos análisis es estudiar los repetidos y secuencias espaciadoras, y determinar su origen.

Los análisis de las secuencias CRISPR, permiten determinar el origen de los espaciadores, así como su localización en el genoma de los bacteriófagos a estudiar. Los *primers* utilizados en estos ensayos fueron diseñados por el equipo del Dr. Sylvain Moineau en la Université Laval, Quebec, Canadá.

Los amplicones obtenidos mediante la amplificación por PCR, fueron secuenciados utilizando los *primers* descritos en las tablas correspondientes (tabla 2). Luego se procedió a ensamblar la secuencia a partir de los *contigs*, utilizando el *software* ApE A plasmid editor (<http://biologylabs.utah.edu/jorgensen/wayned/ape/>). Una vez ensamblada la secuencia, ésta fue analizada haciendo uso del *software* CRISPRfinder (<http://crispr.u-psud.fr>). Por último, las secuencias espaciadoras obtenidas como output, se analizaron mediante BLAST con bases de datos para determinar su origen, o buscar su presencia en genomas virales.

El *software* ApE, permite editar los *contigs* en busca de corregir potenciales errores de secuenciación. Es importante tener especial cuidado en este punto, pues variaciones en una

base en las secuencias espaciadoras pueden llevar a errores importantes. En caso de no poder ensamblar los *contigs*, se debe determinar si el problema es de calidad en la secuencia, o posibles gaps. Si el problema es de calidad, debe reenviarse a secuenciar. Si por el contrario se trata de gaps en la secuencia, deben diseñarse nuevos *primers* para secuenciar las regiones faltantes.

#### **Mapeo de los espaciadores en el genoma de los bacteriófagos**

Para este estudio se utilizaron los bacteriófagos p53, p93, p73, p128 y p107. En la secuencia genómica de cada uno de ellos se ubicaron los espaciadores para todas las cepas analizadas. Se realizó la representación de los genomas para cada *locus* (CRISPR 1 y CRISPR 3) por separado. Para la esquematización se utilizó el software XPlasMap (<http://www.iayork.com/XPlasMap/>).

## **4. RESULTADOS Y DISCUSIÓN**

#### 4.1 Caracterización molecular de *loci* CRISPR en *S. thermophilus*

Con el objetivo de caracterizar molecularmente las cepas de *S. thermophilus*, se realizó el análisis del contenido de los *loci* CRISPR en la colección de cepas comerciales proporcionadas por un proveedor.

La estrategia consistió en la amplificación por PCR de los diferentes *loci* CRISPR de las cepas de la colección y posterior secuenciación y análisis bioinformático. El contenido de espaciadores fue evaluado mediante comparación entre cepas y análisis de bases de datos, a fin de determinar la relación entre ellas, y la potencial resistencia que confiere cada espaciador (bacteriófagos, plásmidos, etc.).

El primer paso para el análisis y caracterización fue la amplificación de los *loci* CRISPR en las cepas de estudio (figura 8). Las cepas de *Streptococcus thermophilus* pueden poseer hasta cuatro *loci* CRISPR diferentes llamados CRISPR 1 (subtipo II---A), CRISPR2 (subtipo III---A), CRISPR 3, (subtipo II---A) y CRISPR4 (subtipo I---E). De estos, generalmente son activos únicamente los del subtipo II---A. Por este motivo se escogieron los *loci* CRISPR 1 y CRISPR 3 para el análisis en este trabajo.

En la figura 8 se presentan a modo de ejemplo los amplicones obtenidos para algunas de las cepas estudiadas. El tamaño de los mismos varía de acuerdo a la cantidad de espaciadores que se hayan incorporado, encontrándose en el rango de entre 1500 y 2000 pares de bases aproximadamente.

De los resultados obtenidos se puede establecer que fue posible amplificar el *locus* CRISPR 1 para todas las cepas (50 cepas) y que CRISPR 3 por su parte también está presente en la gran mayoría, excepto en En7, lo que indicaría que en esta cepa el único *locus* potencialmente activo es CRISPR 1.

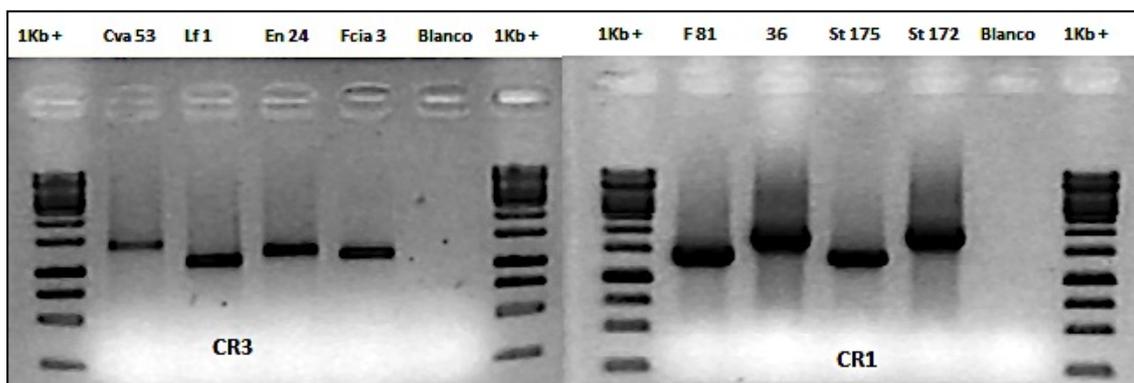


Figura 8. Amplificación de las cepas Cva53, Lf1, En24 y Fcia3 para CRISPR 3 (CR3) y F81, 36, St175 y St172 para CRISPR 1 (CR1). MM: Marcador de peso molecular

#### 4. Resultados y discusión

Se presentan a continuación los resultados del estudio de las secuencias de los *loci* CRISPR 1 y CRISPR 3 de todas las muestras y el análisis de los espaciadores que presentan. Fue posible secuenciar el *locus* completo de 10 cepas para CRISPR 1 y de 32 cepas para CRISPR 3, de las 50 cepas estudiadas y secuenciar parcialmente el *locus* CRISPR 1 de 25 cepas y de 6 cepas para CRISPR 3. No fue posible obtener la secuencia para 15 de las cepas del locus CRISPR 1 y de 12 de las cepas para CRISPR 3, ya sea por errores en el secuenciado, o debido al gran tamaño que podrían presentar algunos de los *locus* que exceden la cobertura del secuenciado (tabla 4). Una de las perspectivas a futuro es el diseño de *primers* internos para amplificar y secuenciar estos fragmentos faltantes.

Tabla 4. Total de cepas secuenciadas. En rosado se marcan aquellas que se ensamblaron completamente, en celeste las incompletas y en verde aquellas que no se ensamblaron. A: para CRISPR 1 y B: para CRISPR 3

**A**

19	11L	Diag 9	EL	En5
En10	F 81	Pot	St179	TH4
CHE102	Cr1	Diag16	En7	En8
En24	Fcia2	Fcia3	G72	Lf1
Lf3	Lf 505	St 2	St43	St62
St172	St175	St309	Lf507	Stsac2
T	TA	Tma	11CH	36
En6	CVA52	CVA 53	Diag12	Hn
3	Lf 2	En10	Lf502	St176
Diag5	St307	Diag1	Diag2	Bio4

**B**

CVA53	Diag9	En4	Fcia3	Hn
Lf1	Pot	St175	St176	St179
T	3	11CH	19	Cr1
Diag5	EL	En 5	En10	F81
Fcia2	G72	Lf2	Lf 3	Lf507
St 307	St309	Stsac2	En8	11L
TH4	Lf502	Diag2	Diag12	En6
36	St2	St62	TA	CHE102
Diag16	Lf505	St43	CVA52	St172
Diag1	Tma	En1	En7	Bio4

#### 4. Resultados y discusión

El análisis se realizó con aquellas cepas en las cuales se logró ensamblar parcial o completamente sus *loci*. A modo ilustrativo se escogió la cepa 19 que posee ambos *loci* secuenciados.

El *locus* CRISPR 1 de la cepa 19 está compuesto por 16 espaciadores y el largo de la secuencia palindrómica que separa los espaciadores es de 36 pares de bases (figura 9). Los espaciadores fueron analizados mediante BLAST, algunos resultados se presentan a continuación (tabla 5). Este estudio también fue realizado para el resto de las cepas, el análisis completo se encuentra en el anexo I para CRISPR 1 y en el anexo II para CRISPR 3.

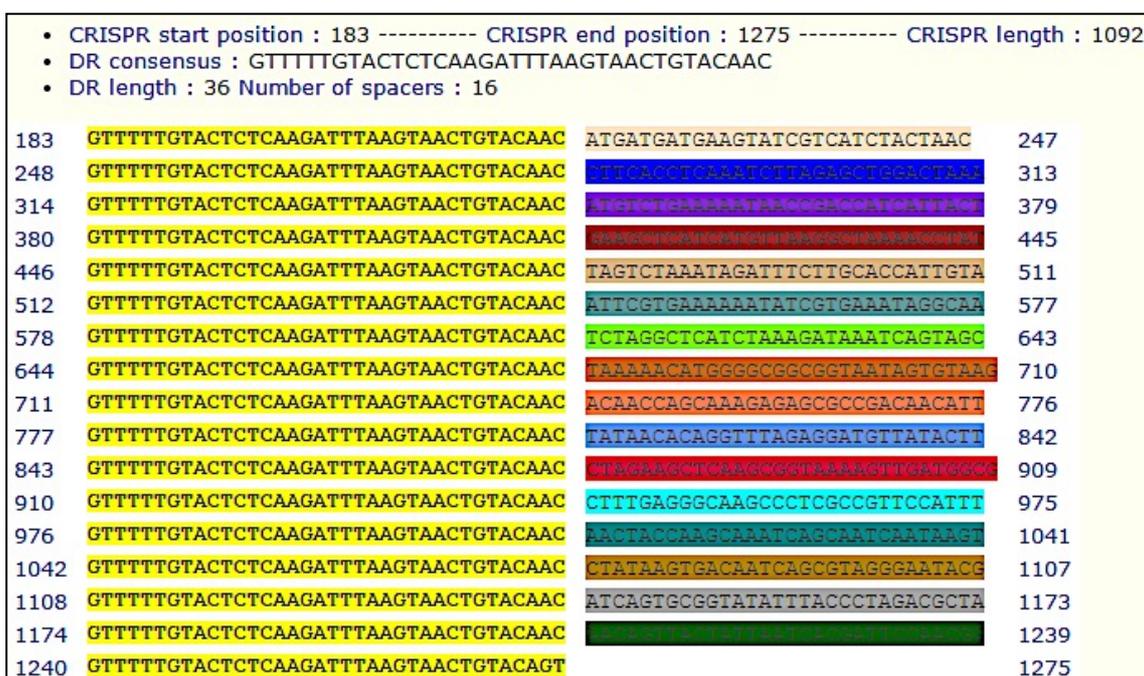


Figura 9. Secuencias palindrómicas y espaciadores del *locus* CRISPR 1 de 19. Esta imagen fue obtenida usando el software CRISPRfinder (<http://crispr.u-psud.fr/>).

Tabla 5. Análisis de los espaciadores del *locus* CRISPR 1 para la cepa 19. en naranja se señalan aquellos espaciadores que únicamente coinciden con secuencias de genomas fágicos y en lila el resultado del BLAST que brindó secuencias de espaciadores de *loci* CRISPR presentes en la base de datos.

1	ATGATGATGAAGTATCGTCATCTACTAAC	9	ACAACCAGCAAAGAGAGCGCCGACAACATT
2	CTTCACCTCAAATCTTAGAGCTGGACTAAA	10	TATAACACAGGTTTAGAGGATGTTATACTT
3	ATGTCTGAAAAATAACCGACCATCATTACT	11	CTAGAAGCTCAAGCGGTA AAAAGTTGATGGCG
4	GAAGCTCATCATGTTAAGGCTAAAACCTAT	12	CTTTGAGGGCAAGCCCTCGCCGTTCCATT
5	TAGTCTAAATAGATTTCTTGACCATTGTA	13	AACTACCAAGCAAATCAGCAATCAATAAGT
6	ATTCGTGAAAAAATATCGTGAAATAGGCAA	14	CTATAAGTGACAATCAGCGTAGGGAATACG
7	TCTAGGCTCATCTAAAGATAAATCAGTAGC	15	ATCAGTGCGGTATATTTACCCTAGACGCTA



## 4. Resultados y discusión

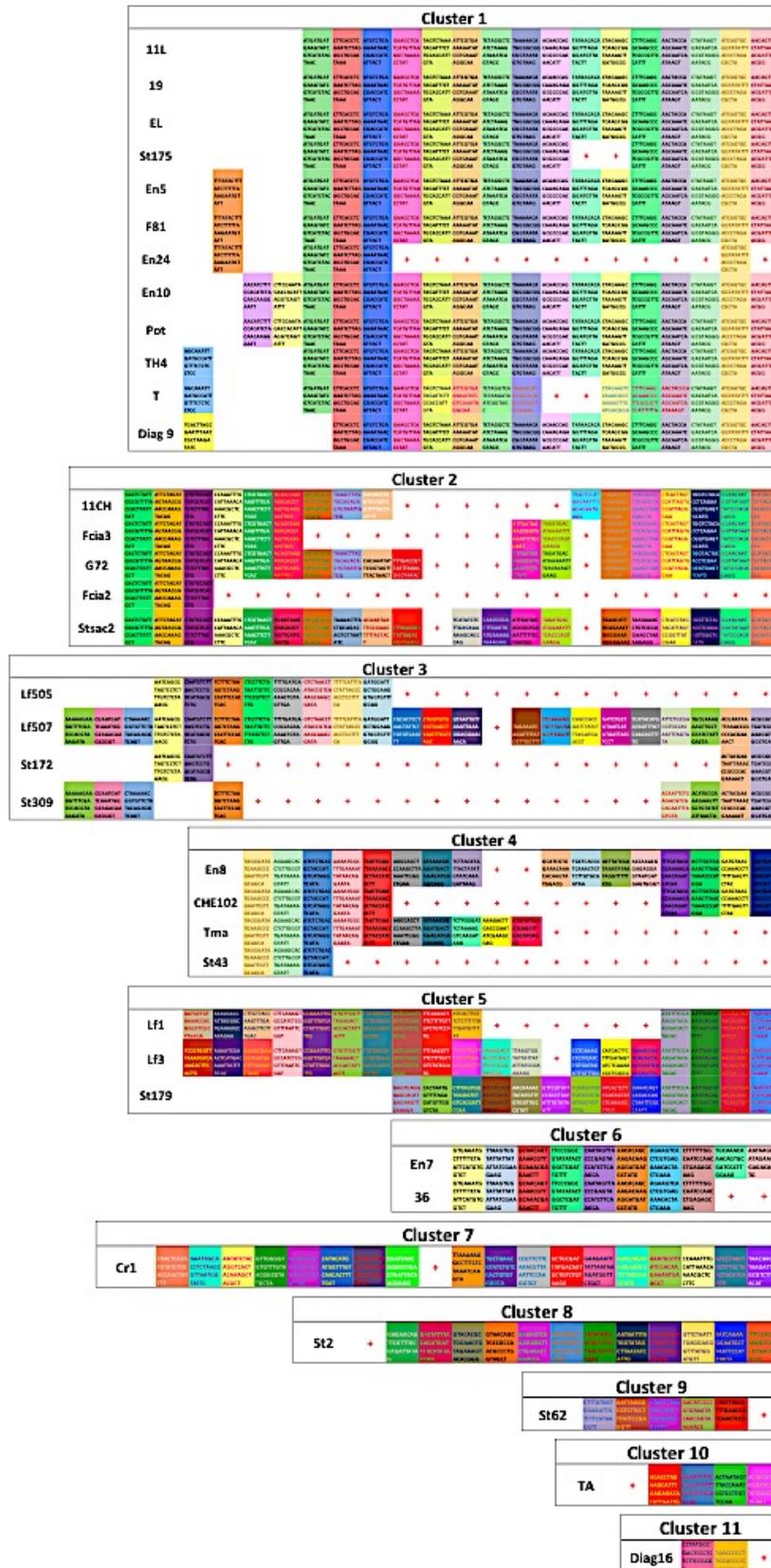


Figura 11. Representación gráfica de la variabilidad de las cepas de *S. thermophilus* utilizando los espaciadores del locus CRISPR 1. Las cepas se agruparon en 11 Clusters. Las secuencias palindrómicas no están incluidas. En cada cluster cada espaciador está representado por una única combinación de color de fondo y texto. Esta combinación permite la exclusiva representación de un espaciador en particular, por lo cual cuadrados con igual combinación de color de fondo y texto representan espaciadores idénticos. Los gaps se esquematizan como cuadrados blanco y con + se representa secuencias incompletas. Los espaciadores se enumeran decrecientemente de izquierda a derecha (5'----3').

La caracterización de los espaciadores del *locus* CRISPR 1 permitió agrupar las cepas de *S. thermophilus*, en 11 *clusters* (figura 11).

El *cluster* 1 está conformado por 12 cepas. A excepción de la cepa En24, cuyo *locus* no fue secuenciado en su totalidad, y basándonos en la teoría de que los espaciadores nuevos se integrarían por el extremo 3' [17] inferiría que todas las cepas de este grupo compartirían el mismo ancestro. El espaciador más antiguo según los datos obtenidos del BLAST corresponden a cuatro diferentes bacteriófagos; *Streptococcus* fago TP-778L, *Streptococcus* fago 858, *Streptococcus thermophilus* bacteriófago 2972, *Streptococcus* fago (anexo I). Dentro de esta agrupación, se observa que determinadas cepas presentan los mismos espaciadores, este es el caso de las cepas 11L, 19, EL y St175, En5 y F81 y por último, las cepas En10 y Pot.

Existen espaciadores que no están presentes en todas las muestras, el espaciador 1 de las cepas En5, F81, En24, En10, Pot, TH4, T, Diag9 y el espaciador 2 de las cepas En10 y Pot (figura 11). Esto puede deberse a que las bacterias por presiones selectivas los hayan perdido o adquirido. Aquellas que no los presentan, puede ser que en su hábitat actual no estén expuestas a esos bacteriófagos y que por tanto no necesiten retener esa información en su genoma. Por el contrario, aquellas que sí los poseen, puede deberse a que en su historia reciente, fueron infectadas por estos bacteriófagos adquiriendo estas nuevas secuencias.

El *cluster* 2 está compuesto por los espaciadores de 5 cepas. El *locus* CRISPR 1 de estas cepas no se pudo secuenciar y ensamblar completamente. Sin embargo podemos observar que la mayoría de estas cepas (salvo Fcia2) podrían compartir el mismo ancestro.

Algunos espaciadores se marcaron con el mismo fondo variando solamente el color de las letras para destacar que son secuencias similares pero no idénticas. Sus diferencias pueden ser consecuencia de mutaciones puntuales, por ejemplo, el espaciador 9 de las cepas 11CH, G72 y Stsac2.

El *cluster* 3 está constituido por cuatro cepas, y al igual que para el *cluster* anterior, no fue posible obtener totalmente secuenciado este *locus*. Exceptuando Lf505 todas las cepas comparten el espaciador más antiguo (figura 11), el cual no está relacionado con ningún fago ingresado hasta el momento en la base de datos.

El *cluster* 4 está integrado también por cuatro cepas, y nuevamente no se logró obtener el *locus* secuenciado en su totalidad. Se puede destacar que todas las cepas comparten el primer espaciador, el cual está relacionado con el genoma del fago 858 (anexo I). Las cepas En8 y CHE102 comparten su último espaciador.

El *cluster* 5 está integrado por 3 cepas. Para las cepas Lf1 y Lf3, tampoco fue posible obtener la secuencia completa. Las tres cepas comparten sus últimos tres espaciadores por lo que

inferimos que tienen una misma historia evolutiva. Su último espaciador está emparentado con los bacteriófagos Sfi 21 y Sfi 19 de origen suizo. En el caso de la cepa Lf3 no fue posible encontrar información relevante en los dos primeros espaciadores, y en el tercer espaciador se encontró relación con el bacteriófago Sfi 18 de origen alemán. Para el primer espaciador de la cepa St179, la base de datos muestra homología con el virus de *Streptococcus* 9874, de origen Irlandés.

El *cluster* 6 está conformado por dos cepas, y solamente se obtuvieron los primeros espaciadores del *locus*. Por ende, no fue posible analizar la pérdida o adquisición de nuevos espaciadores ni conocer si poseían un ancestro en común. Únicamente fue posible inferir que recientemente estuvieron expuestas a los mismos bacteriófagos dado que comparten los primeros espaciadores.

Los últimos cinco *clusters*; 7, 8, 9, 10 y 11 están integrados solamente por una cepa cada uno y sus *locus* están incompletos. Dada la escasa información obtenida, fueron agrupados individualmente ya que no comparten espaciadores con los *clusters* anteriores.

## 4. Resultados y discusión

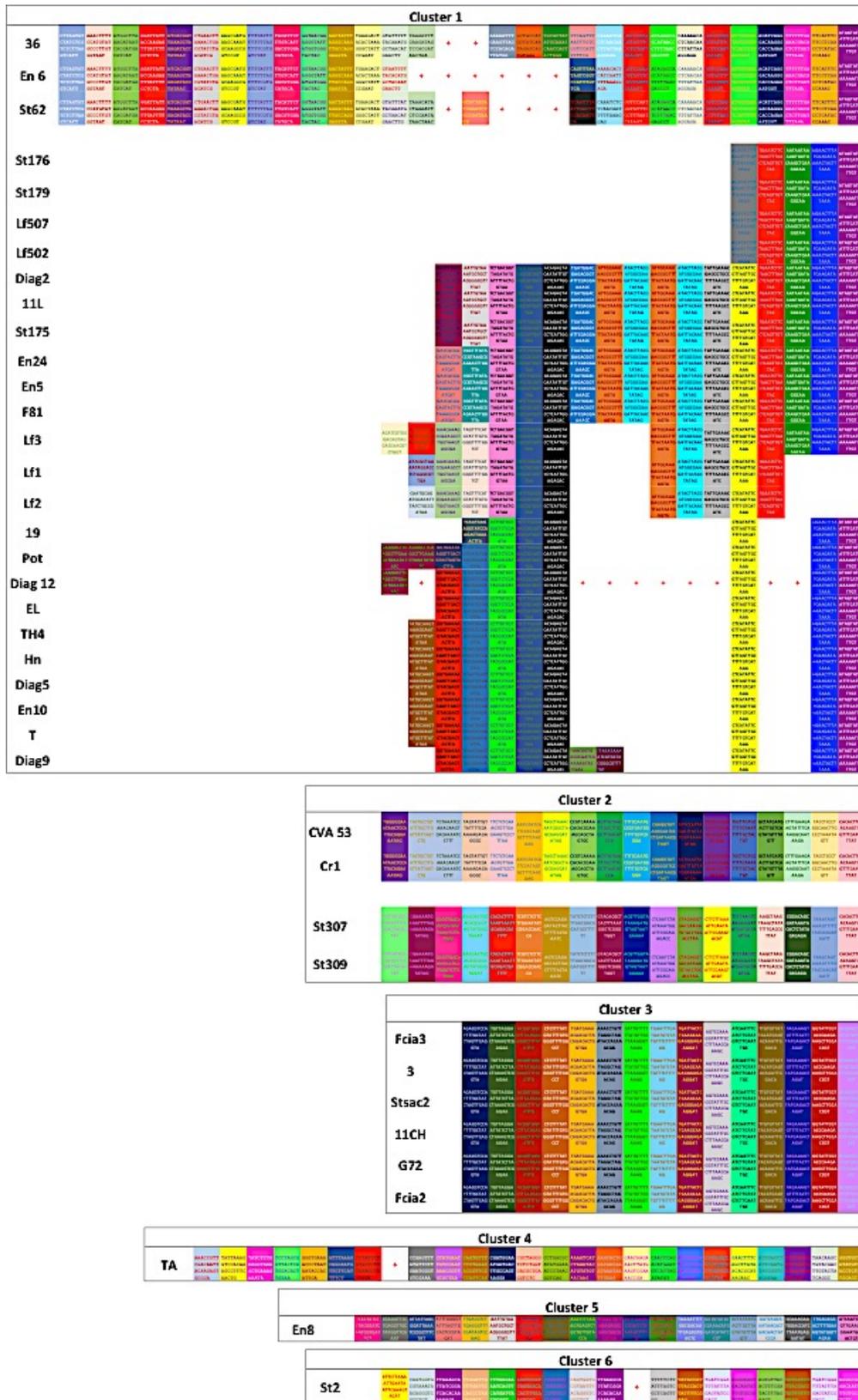


Figura 12. Representación gráfica de la variabilidad de las cepas de *S. thermophilus* utilizando los espaciadores del locus CRISPR 3. Las cepas se agruparon en 6 Clusters. Las secuencias palindrómicas no están incluidas. En cada cluster cada espaciador está representado por una única combinación de color de fondo y texto. Esta combinación permite la exclusiva representación de un espaciador en particular, por lo cual cuadrados con igual combinación de color de fondo y texto representan espaciadores idénticos. Los gaps se esquematizan como cuadrados en blanco y con + se representa secuencias incompletas. Los espaciadores se enumeran decrecientemente de izquierda a derecha (5'-----3').

La caracterización de los espaciadores del *locus* CRISPR 3 agrupó a las cepas de *S. thermophilus* en 6 *clusters* (figura 12).

El *cluster* 1 está compuesto por 25 cepas. Excepto las cepas St175, Lf1 y Lf2, el resto compartirían un mismo ancestro, el cual no se corresponde a ningún fago o plásmido presente hasta el momento en la base de datos consultada (anexo II).

Para las tres primeras cepas (figura 12) no fue posible obtener la secuencia completa del *locus*. Se observan leves cambios en la secuencias de algunos espaciadores, por ejemplo, el espaciador número 14. Estas alteraciones pueden deberse a mutaciones puntuales. Las cepas I) St176, St179, Lf502 y Lf507, II) Diag2, 11L y St175, III) En24, En5, F81, IV) TH4, Hn, Diag5, En10 y T presentan los mismos espaciadores.

Los espaciadores 7 y 8 se hallan duplicados en las cepas Diag2, 11L, St175, En24, En5 y F81. Este duplicado se pierde a partir de la cepa F81, al igual que el espaciador número 6.

Los primeros espaciadores son los que presentan mayor variabilidad. Por ejemplo para las cepas del grupo II los espaciadores 1 y 2 presentan homología con los bacteriófagos TP-J34 y Sfi19 respectivamente (anexo II), y para cepas del grupo III) se observa que el primer espaciador no muestra resultados de homología en la base de datos mientras que el segundo muestra coincidencia con los plásmidos pND103, pt38 y pER16.

El *cluster* 2 está conformado por cuatro cepas, y éstas se pueden agrupar de la siguiente manera; I) St307 y St309 y II) Cva53 y Cr 1.

Los subgrupos I) y II) comparten únicamente el último espaciador, el cual presenta semejanza con la cepa de *S. thermophilus* ND03.

El *cluster* 3 lo componen seis cepas, éstas comparten los mismos 15 espaciadores. Tanto el primer como el último espaciador no muestran resultados de coincidencia con ninguna secuencia de la base de datos. A partir del quinto espaciador se hallan coincidencias en la base de datos, que presentan secuencias semejantes con los bacteriófagos Abc2 y DT1 (anexo II).

Los *clusters* 4, 5 y 6 corresponden a las cepas TA, En8 y St2. Éstas no se agruparon con ninguna otra cepa de los otros *clusters* dado que no poseen ningún espaciador en común.

Para finalizar con el análisis, en la siguiente tabla se detallan aquellas cepas que presentan los mismos espaciadores (o la mayoría de ellos), por un lado las correspondientes al *locus* CRISPR 1 y por otro al *locus* CRISPR 3 (tabla 7).

#### 4. Resultados y discusión

Tabla 7. Cepas con igual *locus* CRISPR 1 y CRISPR 3. \* se refieren a aquellas cepas en el cual su *locus* está incompleto.

		Cepas				
CRISPR 1	11L, 19 y EL	En5 y F81	En10 y Pot	En7 y 36*		
CRISPR 3	St176, St179, Lf507 y Lf502	En24, En5 y F81	TH4, Hn, Diag5, En10 y T	Cva53 y Cr1	St307 y St309	Fcia3, 3, Stsac2, 11CH, G72 y Fcia2

## 4.2 Comparación de los espaciadores encontrados con genomas de bacteriófagos uruguayos

Mediante el análisis previo de los espaciadores de los *loci* CRISPR de las 50 cepas, se procedió a ubicar sus posiciones en los genomas secuenciados de 5 bacteriófagos uruguayos, que pertenecen a una colección mayor de 15 bacteriófagos. Éstos fueron seleccionados, debido a que se contaba con su genoma anotado completamente por nuestro grupo de trabajo [28]. Para los otros 10 bacteriófagos restantes, el ADN fue extraído y enviado a secuenciar a la Universidad Laval de Canadá.

Para este estudio nos planteamos la hipótesis de que aquellos espaciadores que se incorporaron a regiones menos variables del genoma del fago se consideran con mayor potencial para fines biotecnológicos porque serían más estables. Generalmente en un genoma, las regiones menos variables son las que codifican para genes de proteínas no estructurales (lisinas, endonucleasas) y son seguidas en variabilidad por las regiones que codifican para proteínas estructurales. Los espaciadores cuyas secuencias se ubican tanto en regiones intergénicas o que codifican para proteínas hipotéticas, son menos confiables ya que en ellas las mutaciones suelen ocurrir con mayor frecuencia [28].

La evaluación del correcto funcionamiento de los espaciadores, se basará en el criterio de Semenova 2011 [29], los cuales fueron detallados previamente en la introducción,

A continuación se muestra el análisis para la ubicación de los espaciadores del *locus* CRISPR 1 (figura 13-17) y por último los de CRISPR 3 (figura 18-22) en el genoma de los bacteriófagos estudiados.

## 4. Resultados y discusión

### Análisis para el locus CRISPR 1

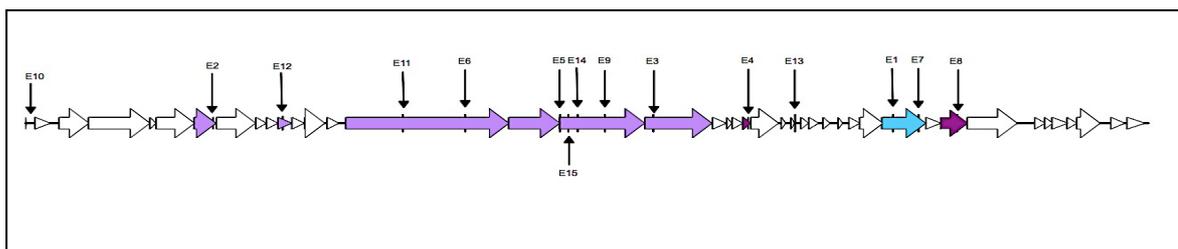


Figura 13. Representación esquemática de los espaciadores del locus CRISPR 1 en el genoma core del bacteriófago de *S. thermophilus* p53. Los espaciadores (E) están numerados del 1 al 15. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 8. Cuadro de referencias de la figura 13. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
GAAGCTCATCATGTTAAGGCTAAAACCTAT	1	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, St175,T	Líder	26437	Si	Si	Si
AACTACCAAGCAAATCAGCAATCAATAAGT	2	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, St175,T	Líder	5706	Si	Si	Si
AACAGTTACTATTAATCACGATTCCAACGG	3	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, St175,T	Líder	19128	Si	Si	Si
CTTTAGTGGTGGGGTCTGTACGATTCCAA	4	St179	Rezagada	22022	Si	No	Si
CCAAATTTGCATTAACAAAACGCTCCTTC	5	11CH, Fcia3, G72, Stsac2, Cr1	Rezagada	16282	Si	Si	Si
TAAGAACTGTGCGAGAAGCTAAAAGAGG	6	Stsac2	Líder	13394	No	Si	Si
TGCTGAACGTATCTGTCCACTGTGTGGCCA	7	Cr1	Líder	27219	No	No	No
AGAAGCACCTCTTGCCTTGATAAAAGTATT	8	En8, CHE102, Tma, St43	Líder	28394	No	No	No
ATCTCTGACGCTACCATGTCATAGGTCATA	9	En8, CHE102, Tma, St43	Rezagada	17655	Si	No	Si
ACCAAACGGAGAGGAGTAATGATGAGTGGGT	10	En8	Líder	272	No	Si	Si
GATGTAACCCAAACCTTTTTGAGTTCTAC	11	En8, CHE102	Rezagada	11494	Si	No	No
AGAAGTCACTCGTGAGAACTACTCTAAA	12	En7, 36	Líder	7828	Si	Si	No
TCATACATGATAGTTTGTCAACACTTTTG	13	Lf507	Rezagada	23450	Si	Si	Si
CTTCAAAGTGCCATCTCGTTTAATTCGGT	14	Lf1, Lf3	Rezagada	16822	Si	Si	Si
ACTAAAATCGTTCTCAATGAGTTGTGACAC	15	Lf1, Lf3	Rezagada	16538	Si	No	Si

En la representación del genoma del bacteriófago p53, se ubicaron 15 espaciadores diferentes del *locus* CRISPR 1 pertenecientes a 25 cepas distintas (tabla 9). Por lo tanto, las cepas que presentarían resistencia ante este fago son las siguientes; 11, 19, Diag9, EL, En5, En10, F81, Pot, St179, TH4, Fcia3, 11CH, G72, Stsac2, Cr1, En8, CHE102, Tma, St43, En7, 36, Lf507, Lf1, Lf3 y T.

De los 15 espaciadores encontrados, E7, E8, y E11 cumplen con las condiciones necesarias, según Semenova [29] para brindar resistencia a las cepas que los adquieran.

A modo de conclusión el *locus* CRISPR 1 le proporcionaría resistencia a las cepas Cr1, En8 y CHE102 de manera más eficiente ante el fago p53 que a las cepas Tma y St43.

## 4. Resultados y discusión

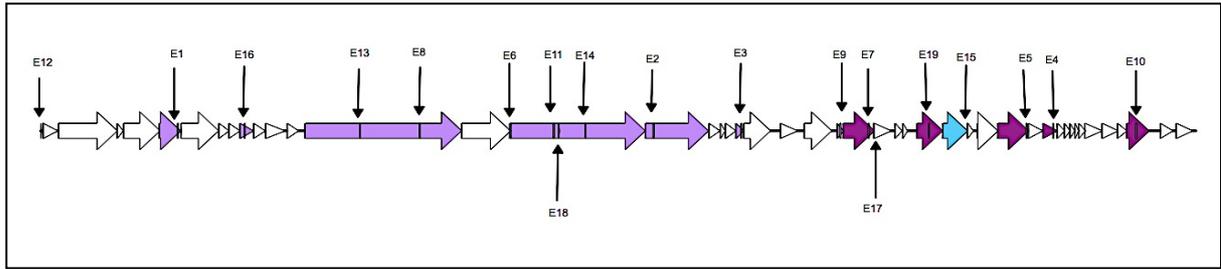


Figura 14. Representación esquemática de los espaciadores del locus CRISPR 1 en el genoma core del bacteriófago de *S.thermophilus* p73. Los espaciadores (E) están numerados del 1 al 19. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 9. Cuadro de referencias de la figura 14. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
AACTACCAAGCAAATCAGCAATCAATAAGT	1	11L,19,Diag9, EL,En5,En10,F81, Pot,TH4,St175	Líder	4351	Si	Si	Si
AACAGTACTATTAATCAGGATTCCAACGG	2	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, St175, T	Líder	19477	No	Si	Si
CTTTAGTGGTGGGGTCTGTCCAGATTCCAA	3	St179	Rezagada	22267	No	Si	Si
ATCACTCTTTGATAGTATCTCAAACGCTGG	4	St179	Líder	32185	Si	No	Sí
AGCAAATTGATGCCATTGTTTCTCTCTCC	5	St179	Rezagada	31335	Si	No	Si
CCAAATTTGCATTAACAAAACGCTCCTTC	6	11CH,Fcia3,G72, Stsac2,Cr1	Rezagada	14917	Si	Si	Si
CTCGTAACTAAGTTTGAAACTTCTTTGAC	7	11CH,Fcia3, G72, Stsac2, Lf1	Líder	26318	Si	No	Si
TAAGAAACCTGTCGGAGAAGCTAAAAGCTG	8	11CH,Fcia3, G72, Stsac2	Líder	12039	No	Si	Si
TGGTACTACACTCGAACGGTGAGTTCATG	9	G72, Stsac2	Líder	25416	Si	Si	No
TCAGCATTGGTTTACATGACCCACGCTG	10	Cr1	Rezagada	34826	Si	Si	No
ATCTCTGACGCTACCATGTCATAGGTCATA	11	En8,CHE102, Tma, St43	Rezagada	16300	Si	No	Si
ACCAAACGGAGAGGAGTAATGATGAGTGGGT	12	En8	Líder	62	No	Si	Si
GATGTAACCCAACCTTTTTTGAGTTCTAC	13	En8, CHE102	Rezagada	10138	Si	No	No
AAATGACGGAACGTATAAAGGTCTAACCGT	14	En8, CHE102	Líder	17308	Si	Si	Si
AACACAGCAAGACAAGAGGATGATGCTATG	15	En7, 36	Líder	29451	Si	Si	No

#### 4. Resultados y discusión

AGAACTCACTCGTGAGAACTACTCAA	16	En7, 36, Lf1, Lf3	Líder	6473	Si	Si	No
CTCCTTCTATAATTGTTCTCCGTTCTCG	17	Lf505, St507	Rezagada	26475	Si	No	Si
GTGTTGGTTAAAACTACCACTATTACTT	18	Lf1, Lf3	Líder	16473	Si	No	Si
GTACACGCTGGAAGTGTAGAAAGTACACGGG	19	St2	Líder	28232	No	Si	Si

En el esquema genómico del fago p73, se ubicaron 19 espaciadores distintos del *locus* CRISPR 1 pertenecientes a 28 cepas (tabla 10). Las siguientes cepas, son posibles candidatas a presentar resistencia; 11L, 19, Diag9, EL, En5, En10, F81, Pot, St179, TH4, 11CH, Fcia3, G72, Stsac2, Cr1, En8, CHE 102, Tma, St43, En7, 36, Lf505, St507, St175, Lf1, Lf3, St2 y T.

De los 19 espaciadores obtenidos E4, E5, E7, E10, E17 y E19 se ubican en regiones correspondientes a proteínas hipotéticas y E9 y E12 están ubicados en regiones intergénicas. Contra el fago p73 el *locus* CRISPR 1, solamente cuenta con el espaciador E13 capaz de brindar resistencia, y este lo contiene únicamente la cepa En8.

## 4. Resultados y discusión

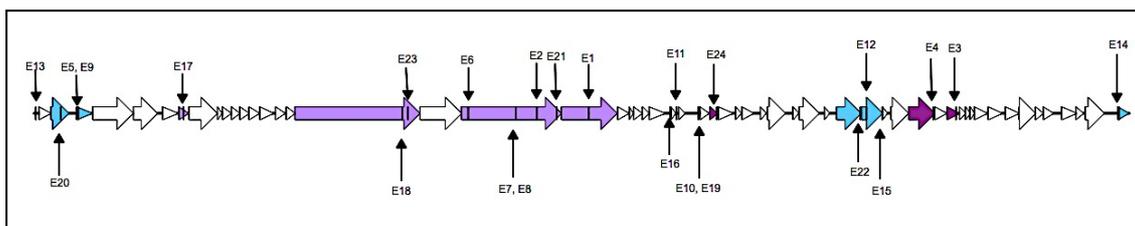


Figura 15. Representación esquemática de los espaciadores del locus CRISPR 1 en el genoma core del bacteriófago de *S. thermophilus* p93. Los espaciadores (E) están numerados del 1 al 24. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 10. Cuadro de referencias de la figura 15. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables

Espaciador	Nº de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
CTTCACCTCAAATCTTAGAGCTGGACTAAA	1	11L,19,Diag9, EL,En5, En10, F81, Pot, TH4, En24	Rezagada	20530	Si	Si	No
TAGTCTAAATAGATTTCTGCACCATTGTA	2	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4	Rezagada	18611	Si	No	Si
ATCACTCTTTGATAGTATCTCAAACGCTGG	3	St179	Líder	34132	Si	No	Si
AGCAAATTGATGCCATTGTTCTCTCTCC	4	TH4	Rezagada	33288	No	No	Si
ACAGTAAGATACACGTAGTTGATGAATTGG	5	11CH, Fcia3, G72	Líder	1625	Si	Si	Si
AAGACATATCTTTTTAACATCATGAAGACG	6	11CH, G72	Líder	16064	Si	Si	No
ATTGATAGAACCGGTGAAATTTTGCCAAT	7	Fcia3, Stsac2	Líder	17828	No	Si	No
ATTGATAGACCGGTGAAATTTTCTATC	8	G72	Líder	17828	Si	Si	No
ACAGTAAGATACACGTAGTTGATGAATTG	9	Stsac2	Líder	1625	No	Si	Si
CATACATGATAGTTTGCAACACTTTTGAT	10	Cr1, Tma	Rezagada	24592	Si	Si	No
AGGCAGAAAAGAAGTATTTGGTAAGTATG	11	Cr1	Líder	23779	Si	Si	No
TGCTCACGCTCAAAGTCTTGTGCGATAA	12	En8	Rezagada	30821	Si	No	Si
TTAAGTGGTATTATTATATTATCGAAGAAG	13	En7, 36	Líder	47	No	No	No
GCAACAGTAAAACGTTGCAAACGAAACTT	14	En7, 36	Líder	40174	Si	Si	No
AACACAGCAAGACAAGAGGATGATGCTATG	15	En7, 36	Líder	31404	Si	Si	Si
TCTTTCTAAAGTCTAAGCACTTCCACTCAC	16	Lf505, Lf507, St309,	Rezagada	23553	Si	No	Si
TTTTGATGACCGCACAAAACTGATGTTGA	17	Lf505, Lf507	Líder	5524	No	No	Si
CGAATGATTCAAATGGGCAGACAAGCCGGT	18	Lf507, St309	Líder	13622	No	No	Si
TCATACATGATAGTTTGCAACACTTTTG	19	Lf507	Rezagada	24594	Si	Si	Si
AATGCTGTAAAACCCGCGCTCGCTTAAGA	20	Lf1	Rezagada	969	No	No	No
ATACGCCATTAGGACTGAAAAGCATAGAA	21	Lf1	Rezagada	19348	Si	No	No
CCGAATTTGGGTTTATCACTATTTGGCTTG	22	Lf1, Lf3	Rezagada	30590	Si	Si	No
GTAACAGCTGCCCAACGCCCTGGTAGCT	23	St2	Líder	13834	Si	Si	Si
AATAATTTATGGTATAGCTTAATATCATTG	24	St2	Líder	25273	Si	No	Si

En el genoma del fago p93 se ubicaron 24 espaciadores, distribuidos en las siguientes 26 cepas; 11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, 11CH, Fcia3, G72, Stsac2, Cr1, En8, Tma, En7, 36, Lf505, Lf507, St309, En24, Lf1, Lf3, St2.

De los 24 espaciadores E3, E4 y E24 se sitúan en regiones que codifican para proteínas hipotéticas y E10, E11, E13, E16 y E19, en regiones intergénicas.

Solamente los espaciadores E20 y E21 cumplirían con las condiciones óptimas para otorgarle resistencia a las cepas que los contengan. Por lo tanto la cepa Lf1 va a ser la única que presentaría resistencia en su *locus* CRISPR 1 para el bacteriófago p93.

## 4. Resultados y discusión

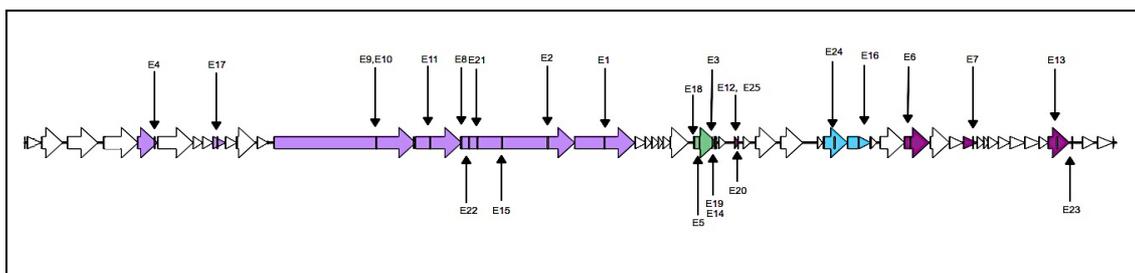


Figura 16. Representación esquemática de los espaciadores del locus CRISPR 1 en el genoma core del bacteriófago de *S. thermophilus* p107. Los espaciadores (E) están numerados del 1 al 25. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 11. Cuadro de referencias de la figura 16. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
CTTCACCTCAAATCTTAGAGCTGGACTAAA	1	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, En24,T	Rezagada	19608	Si	Si	No
TAGTCTAAATAGATTTCTTGCACCATTTGTA	2	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4	Rezagada	17689	Si	Si	Si
TCTAGGCTCATCTAAAGATAAATCAGTAGC	3	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4	Líder	23336	Si	Si	Si
AACTACCAAGCAAATCAGCAATCAATAAGT	4	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4	Líder	4384	Si	Si	Si
AACATCTTTCCACATCTACAACAAGGAATT	5	En10, Pot	Rezagada	22826	Si	Si	No
GAGTGAGGAAGCACATGAGAAGTTCAAAGA	6	St179	Líder	29966	No	No	Si
ATCACTCTTTGATAGTATCTCAAACGCTGG	7	St179	Líder	32088	Si	Si	Si
CCAAATTTGCATTAACAAACAAACGCTCCTTC	8	11CH, Fcia3, G72, Stsac2, Cr1	Rezagada	14767	Si	Si	Si
TAAGAAACCTGTCGGAGAAGCTAAAAGCGG	9	11CH, Fcia3, G72,	Líder	11882	No	Si	Si
TAAGAAACCTGTCGGAGAAGCTAAAAGAGG	10	Stsac2	Líder	11882	No	Si	Si
AATATCTACAGGTCACTACAAAGCTACGCT	11	Cr1	Líder	13705	Si	Si	No
CATACATGATAGTTTGTCAACACTTTTGAT	12	Cr1	Rezagada	24129	No	No	Si
TCAGCATTGGTTTACATGACCCACGCTCTG	13	Cr1	Rezagada	34930	Si	Si	Si
AGGCAGAAAAGAAGTATTTGGTAAGTATG	14	Cr1	Líder	23383	Si	No	No
ATCTCTGACGCTACCATGTCATAGGTCATA	15	En8, CHE102, Tma, St43	Líder	16140	Si	No	Si
AACACAGCAAGACAAGAGGATGATGCTATG	16	En7,36	Líder	28213	No	No	Si
AGAAGTCACTCGTGAGAAACACTACTCAAA	17	En7,36	Líder	6506	Si	Si	No
TCTTTCTAAAGTCTAAGCACTTCCACTCAC	18	Lf505,Lf507, St309	Rezagada	22629	Si	No	Si
CAGCCACCGATTTATCTTTAGATGAACCT	19	Lf507	Rezagada	23339	No	Si	Si
TCATACATGATAGTTTGTCAACACTTTTG	20	Lf507	Rezagada	24131	No	No	Si

#### 4. Resultados y discusión

CTTCAAAGTGCCATCTCGGTTAATTCGGT	21	Lf1, Lf3	Rezagada	15307	Si	Si	Si
ACTAAAATCGTTCTCAATGAGTTGTGACAC	22	Lf1, Lf3	Rezagada	15022	Si	No	Si
AAAGCGAGGCTGGTGGCTTTGAAATCTAAC	23	Lf3	Líder	35440	Si	Si	Si
GTACACGCTGGAAGGTAGAAAAGTACACGGG	24	St2	Líder	27397	No	Si	Si
CATATAAGATAGTTTGTCAACACTTACGA	25	St2	Rezagada	24129	No	No	Si

Se ubicaron en el genoma del fago p107 25 espaciadores, contenidos en el *locus* CRISPR 1 de las siguientes 29 cepas; 11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, 11CH, Fcia3, G72, Stsac2, Cr1, En8, CHE102, Tma, St43, En7, 30, Lf505, Lf507, St309, En24, Lf1, Lf3, St2 y T.

Del total de los 25 espaciadores E6, E7, E12, E13, E20 y E25 se encuentran en regiones que codifican para proteínas hipotéticas y E23 se encuentra en una región intergénica.

En el caso de este fago solamente la cepa Cr1 que presenta el espaciador E14, adquiriría resistencia, ya que este espaciador es el único que cumple con los requisitos necesarios.

Por lo tanto el espaciador 14 contenido en el *locus* CRISPR 1 es el único capaz de conferir resistencia ante el fago p107.

## 4. Resultados y discusión

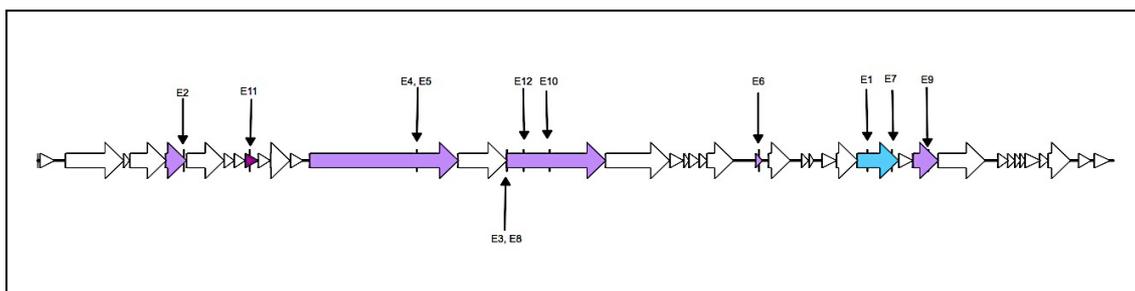


Figura 17. Representación esquemática de los espaciadores del locus CRISPR 1 en el genoma core del bacteriófago de *S. thermophilus* p128. Los espaciadores (E) están numerados del 1 al 19. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 12. Cuadro de referencias de la figura 17. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
GAAGCTCATCATGTTAAGGCTAAACCTAT	1	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, En24,T	Líder	26674	No	No	Si
AACTACCAAGCAAATCAGCAATCAATAAGT	2	11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4	Líder	4686	Si	Si	Si
CCAAATTTGCATTAACAAAACGCTCCTTC	3	11CH, Fcia3, G72, Stsac2	Rezagada	15077	Si	Si	Si
TAAGAAACCTGTCGAGAGCTAAAAGCGG	4	11CH, Fcia3, G72	Líder	12183	No	Si	Si
TAAGAAACCTGTCGAGAGCTAAAAGAGG	5	Stsac2	Líder	12183	No	Si	Si
TGGTCTACACCTCAAACGGTGAGTGCATG	6	Stsac2	Líder	23186	Si	No	No
TGCTGAACGTATCTGTCCACTGTGTGGCCA	7	Cr1	Líder	27456	Si	No	Si
CCAAATTTGCATTAACAAAACGCTCCTTC	8	Cr1	Rezagada	15078	Si	No	No
AGAAGCACCTCTTGCCTTGATAAAAGTATT	9	En8, CHE102, Tma, St43	Líder	28631	Si	No	No
ATCTCTGACGCTACCATGTCATAGGCATA	10	En8, CHE102, Tma, St43	Rezagada	16450	No	No	Si
AGAAGTCACTCGTGAGAACTACTCAAA	11	En7,36	Líder	6808	Si	Si	No
CTTCAAAGTGCCATCTCGGTTTAATTCGGT	12	Lf1, Lf3	Rezagada	15617	Si	Si	Si

En la representación del genoma de p128 se localizaron 12 espaciadores diferentes, presentes en 24 cepas, que se nombran a continuación; 11L, 19, Diag9, EL, En5, En10, F81, Pot, TH4, 11CH, Fcia3, G72, Stsac2, Cr1, En8, CHE102, Tma, St43, En7, 36, En24, Lf1, Lf3 y T.

Del total de los 11 espaciadores E12 se sitúa en una región que codifica para proteínas hipotéticas.

Se encontraron 3 espaciadores, E6, E8 y E9, que cumplen con todos los requisitos según Semenova [29] para otorgarle resistencia a las cepas que los contengan. Por lo tanto las cepas StSac2, Cr1, En8, CHE102, Tma y St43 serían resistentes al fago p128, debido a que tienen algunos de estos espaciadores en su *locus* CRISPR 1.

## 4. Resultados y discusión

### Análisis para el locus CRISPR 3

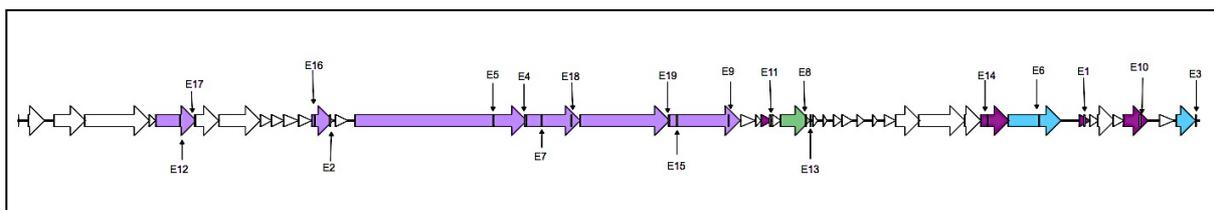


Figura 18: Representación esquemática de los espaciadores del locus CRISPR 3 en el genoma core del bacteriófago de *S. thermophilus* p53. Los espaciadores (E) están numerados del 1 al 18. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 13. Cuadro de referencias de la figura 18. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
<b>TCTAAATCCAACAAGTCACAAAATCCTTT</b>	1	Cva53, Cr1	Rezagada	30906	No	No	No
TAACGGTGAGCCAATCAAAAAATACTTCAC	2	Diag9	Líder	9020	Si	No	Si
TTATATAAAATCGTGCTACGGGCGTTTTAT	3	Diag9	Rezagada	34142	Si	Si	No
<b>CTCATATTCGTTAGTTGCTTTTGCATAAA</b>	4	Diag9, En24, Hn, Lf1, Pot, St175, T, 19, Diag5, EL, En5, En10, F81, Lf2, Lf3, 11L, TH4, Daig12, Diag2	Rezagada	14679	Si	No	No
ATACTTACGATGGCGAAGATTACAACATATAG	5	En24, Lf1, St175, En5, F81, Lf2, Lf3, 11L, Diag2	Líder	13759	Si	Si	No
TGAATCTTCTAACTTTAACTCAGTTGTTAC	6	En24, Lf1, St175, St176, St179, En5, F81, Lf2, Lf3, Diag2, Lf507, 11L, Lf502	Líder	29589	Si	Si	No
<b>AAACAAAGCCGAAGCCTTGTAAGTAGCGA</b>	7	Lf1, Lf2, Lf3	Líder	15160	No	No	No
TGATGAAAACAACATTACAGACTGGTGA	8	Fcia3, 3, 11CH, Fcia2, G72, StSac2	Líder	22825	No	Si	No
TATCTCTCTTTTAACGACAGATGGTTTTTTT	9	St307, St309	Rezagada	20584	Si	Si	No
CGGACAGCGATAAATACACTCTATACAGAGA	10	St307, St309	Líder	32507	Si	No	Si
CCAAAGAATGGACCATCTTAATGAGAATAT	11	En8	Líder	21784	Si	Si	Si
AAGCGATGAAGCAAATGCAAGGCGGTCCTGT	12	36, En6, St62	Líder	4673	No	No	Si
TTTTATTTTTTTTTAGTTTTCGTGGTCTAT	13	36, En6, St62	Rezagada	22960	Si	No	Si
<b>TACATTCTTTGTCAITGACGTGGACATGCA</b>	14	36, En6, St62	Líder	28097	No	No	No
TTGGATTGTTTGATTCCACCTTAGCACACT	15	St2	Líder	19087	Si	Si	No
<b>TATTAAGGTCCACAAAGCCTTCAACTG</b>	16	TA	Líder	8582	Si	No	No
<b>GGGTGAAAACCCTAATAATACCACATTGA</b>	17	TA	Líder	5097	Si	No	No

#### 4. Resultados y discusión

AAAATCATCTAAATACATGTGTGTAACAAG	18	TA	Rezagada	16022	No	Si	No
--------------------------------	----	----	----------	-------	----	----	----

En la representación del genoma del bacteriófago p53, se lograron ubicar 19 espaciadores del *locus* CRISPR 3, pertenecientes a las siguientes 38 cepas; Cva53, Diag9, En24, Fcia3, Hn, Lf1, Pot, St175, St176, St179, T, 3, 11CH, 19, Cr1, Diag5, EL, En5, En10, F81, Fcia2, G72, Lf2, Lf3, Lf507, St309, Stsac2, En8, 11L, Diag2, TH4, 36, En6, St62, Diag12, Lf502, St2, y TA.

Del total de los 18 espaciadores E1, E10, E11, E14, se ubicaron en regiones que codifican para proteínas hipotéticas. Los espaciadores E4, E7, E16 y E17 se ubican en regiones conservadas que codifican para proteínas estructurales y a la vez cumplen con todas las condiciones para brindarle resistencia a las cepas que los contengan.

En conclusión, las cepas Cva53, Cr1, 36, En6 y St62 presentarían resistencia en su *locus* CRISPR 3 aunque de manera menos eficientes que las cepas Diag9, En24, Hn, Lf1, Pot, St175, T, 19, Diag5, EL, En5, En10, F81, Lf2, Lf3, 11L, Diag2, TH4, Daig12, Lf1 y TA, ya que éstos contienen los espaciadores situados en la regiones conservadas. Además, la cepa TA sería la que presente mejor resistencia en el *locus* CRISPR 3 ante el fago p53 ya que contiene dos de los espaciadores que se encuentran en la región más conservada del genoma.

## 4. Resultados y discusión

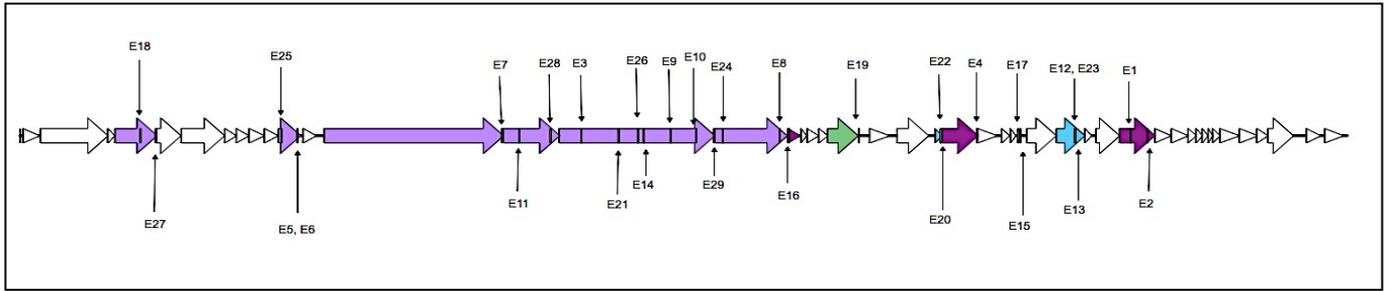


Figura 19. Representación esquemática de los espaciadores del locus CRISPR 3 en el genoma *core* del bacteriófago de *S.thermophilus* p73. Los espaciadores (E) están numerados del 1 al 30. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 14: Cuadro de referencias de la figura 19. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables.

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
<b>TTCTCTCAAAGTGTGAGAAGTCCCTTTAA</b>	1	Cva53, Cr1	Rezagada	30729	No	No	No
TAGCTAAACAATCGGCTAGCGAGGATATAG	2	Cva53, Cr1	Líder	31195	No	Si	Si
TTTCCAATGCCGTGATGGTTTTCGTCAGGA	3	Cva53, Cr1	Rezagada	15523	Si	No	Si
<b>TATCCAGAGAATGGAAGAACAATTATAGA</b>	4	Cva53, Cr1	Líder	26470	Si	No	No
TAACGGTGAGCCAATCAAAAACTTCAC	5	Diag9	Líder	7665	Si	No	Si
TTATATAAATCGTGCTACGGGCGTTTTAT	6	Diag9	Rezagada	36640	Si	Si	No
CTCATATTCGTTAGTTGCTTTTGTCAAAA	7	Diag9, En24, Hn, Lf1, Pot, St175, T, 19, Diag5, EL, En5, En10, F81, Lf2, Lf3,11L,Diag2, TH4 , Diag12	Rezagada	13324	Si	Si	No
TCTGACGGTTAGATATGATTTTACTGGTAA	8	En24, Lf1,St175, En5, F81, Lf2, Lf3, 11L, Diag2	Líder	21013	No	Si	No
TGATGGACGAGACGGTATCCAGGAAAACC	9	En24, St175, En5, F81, 11L, Diag2	Líder	17995	Si	Si	No
<b>ATCAATTTTCATCTTGTATGTCTTCAATTGC</b>	10	Fcia3, 3, 11CH, Fcia2, G72, Stsac2,	Rezagada	18701	Si	No	No
<b>AAACAAAGCCGAAGCCTTGGTAAGTAGCGA</b>	11	Lf1, Lf2, Lf3	Líder	13805	No	No	No
GAGGACTTCAGCCTTGAAGCTGAATATATC	12	Pot, Diag12	Líder	29191	Si	No	Si
AGGACTTCAGCCTTGAAGCTGAATATATC	13	Pot	Líder	29192	Si	No	Si
GAATTTGCTTGAAGGGACTAAAGACTTTAG	14	St175, 11L, Diag2	Líder	17248	No	Si	No
<b>CTTCTAAAATTGAATAATTCGAAGTACAT</b>	15	St307, St309, St2	Líder	27673	Si	No	No
TTCTAAGTCAAAGAATATGATGCTAGTGG	16	St307, St309	Líder	21274	Si	No	Si
ATTTGGGGTATTAGTTGCACTCCGTACAT	17	En8	Rezagada	27612	Si	Si	No
AAGCGATGAAGCAAATGCAAGGCGTCCGT	18	36, En6, St62	Líder	3318	No	No	Si
TTTTATTTTTTTAGTTTTCGTGGTCTAT	19	36, En6, St62	Rezagada	23205	Si	Si	Si
<b>GTATTTTTTACAAATGGCTAACATGAACTT</b>	20	36, En6	Rezagada	25474	Si	No	No
GGTCAAAGGTATTTCCCTTAATTAAGT	21	36, En6	Rezagada	16559	Si	Si	No

#### 4. Resultados y discusión

GTATTTATTACAAATGGCTAACATGAACTTG	22	St62	Rezagada	25474	Si	Si	No
GAGGACTTCAGCCTTGAAGCTGAATATATC	23	Pot, Diag12	Líder	29191	Si	No	Si
TTGGATTGTTGATTCCACCCTAGCACACT	24	St2	Rezagada	19437	Si	Si	No
TATTAAGGTCACAAAGCCTTCAACTG	25	TA	Líder	7227	Si	No	No
TCCTAGCGATTACTCATGGACACTTGTA	26	TA	Líder	17092	Si	Si	No
GGGTGAAAACCTAATAATACCACATTGA	27	TA	Líder	3742	Si	No	No
AAAATCATCTAAATACATGTGTGTAACAAG	28	TA	Rezagada	14667	No	Si	No
CAACCCAGACATGAATGTCATTAGATATGT	29	TA	Líder	19192	Si	Si	No

En el genoma del fago p73 se hallaron 30 espaciadores, que se encuentran en las 34 cepas siguientes; Cva53, Diag9, En24, Fcia3, Hn, Lf1, Pot, St175, T, 3, 11CH, 19, Cr1, Diag5, EL, En5, En10, F81, Fcia2, G72, Lf2, Lf3, St309, Stsac2, En8, 11L, Diag2, TH4, 36, En6, St62, Diag12, St2 y TA.

De los 30 espaciadores E1, E2, E4 E16 y E20 se encuentran en regiones que codifican para proteínas hipotéticas y E15 y E17 en regiones intergénicas.

Por otro lado los espaciadores E10, E11, E25 y E27 están situados en regiones conservadas, las cuales codifican para proteínas estructurales y a su vez no tienen mutaciones en sitios relevantes, por ende, aquellas cepas que los adquirieran contarían con una mejor resistencia en su *locus* CRISPR 3 ante el fago p73. Las cepas son las siguientes; Fcia3, 3, 11CH, Fcia2, G72, Stsac2, Lf1, Lf2, Lf3 y TA.

Finalmente, la cepa TA sería la que presente mejor resistencia en el *locus* CRISPR 3 ante el fago p73 ya que contiene dos de los espaciadores que se encuentran en la región más conservada del genoma, mientras las demás cepas solamente presentan uno u otros espaciadores ubicados en regiones menos conservadas.

## 4. Resultados y discusión

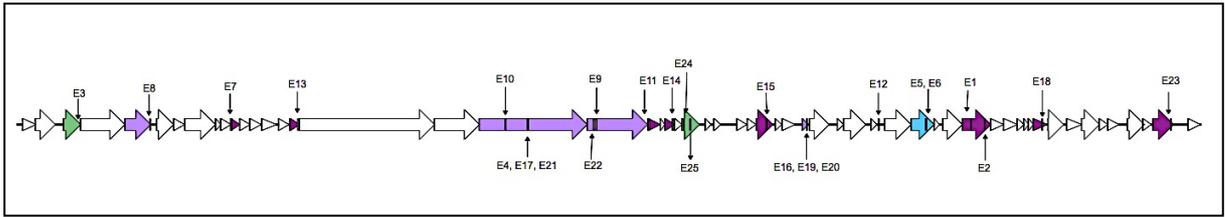


Figura 20. Representación esquemática de los espaciadores del locus CRISPR 3 en el genoma core del bacteriófago de *S. thermophilus* p93. Los espaciadores (E) están numerados del 1 al 25. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 15. Cuadro de referencias de la figura 20. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
<b>TTCTCTCAAAGTGTGAGAAGTCCCTTAA</b>	1	Cva53, Cr1	Rezagada	32683	No	No	No
TAGCTAAACAATCGGCTAGCGAGGATATAG	2	Cva53, Cr1	Líder	33148	No	Si	Si
GCTATGATGACTTGCTCAGTATGTTAGTT	3	Cva53, Cr1, En8	Líder	2138	Si	Si	No
AGTCCAAAGGTATTTCCCTTAAACCAAAGC	4	Fcia3, 3, 11CH, Fcia2, G72, Stsac2	Rezagada	17489	Si	No	Si
GAGGACTTCAGCCTTGAAGCTGAATATATC	5	Pot, Diag12	Líder	31144	Si	Si	Si
AGGACTTCAGCCTTGAAGCTGAATATATC	6	Pot, Diag12	Líder	31145	Si	Si	Si
TCTTATCGGTATATTTAGCACTACGGTAT	7	St307, St309	Rezagada	7326	No	Si	Si
<b>AACGACTGCACAGCGGCAGTATGGGTGAAT</b>	8	St307, St309	Rezagada	4532	Si	No	No
<b>GTACACGCTCACTTTAATGGGTGCGGTGGT</b>	9	St307, St309	Líder	19828	No	No	No
CTCAATCTAATAAATGGATTCGAAAGACC	10	St307, St309	Líder	16716	No	Si	Si
TTCTAAGTCAAAGAATATGATGCTAGTGG	11	St307, St309	Líder	21642	Si	No	Si
ATTTGGGTATTTAGTTGCACTCCGTACAT	12	En8	Rezagada	29523	Si	Si	No
<b>TATTCGAGCTCGCTCAACAGTTTTTTGT</b>	13	En8	Rezagada	9601	No	No	No
CCAAAGAATGGACCATCTTAATGAGAATAT	14	En8	Líder	22440	Si	Si	No
ATGGCTTAAAGATAATCACGATGACATCAT	15	36, En6, St62	Líder	2566	Si	Si	No
<b>GTATTTTTACAAATGGCTAACATGAACCT</b>	16	36	Rezagada	27054	No	No	No
GGTCCAAAGGTATTTCCCTTAAATTAAGT	17	36, En6, St62	Rezagada	17489	Si	Si	Si
TTCATTTCTCTCAACCTCATACCCAAC	18	36, En6, St62	Rezagada	35122	Si	Si	No
GTAATTTTTACACATGGCTACAATGAACCT	19	En6	Rezagada	27054	Si	Si	No
GTATTTATTACAAATGGCTAACATGAACCTG	20	St62	Rezagada	27054	Si	Si	No
GGTCCAATGGTATTTCCCTTAAATTAAGT	21	St62	Rezagada	17489	Si	Si	Si
<b>TTGGATTGTTGATTCCACCCTAGCACACT</b>	22	St2	Rezagada	19741	Si	No	No
TGATCGGATTTATTTAGAGTATCGCACC	23	St2	Líder	39542	Si	Si	No

#### 4. Resultados y discusión

ATTGACCTATTCAATGTATGGGTCACGTAA	24	TA	Líder	22837	No	Si	Si
GGGTGCTGATATGTCTACCTCTGGTGGAGC	25	TA	Líder	23052	No	Si	Si

Se localizaron en el genoma del fago p93 25 espaciadores distribuidos en las siguientes 18 cepas: Cva53, Fcia3, pot, 3, 11CH, Cr 1, Fcia2, G72, St307, St309, Stsac2, En8, 36, En6, St62, Diag12, St2 y TA.

Los E1, E2, E7, E11, E13, E14, E15, E18 y E23 se localizaron en regiones que codifican para proteínas hipotéticas y E12 en una región intergénica.

Por otro lado los espaciadores E8, E9, E16 y E22 se sitúan en regiones conservadas del genoma y a su vez son capaces de brindarle resistencia a las cepas que los porten, que son: St307, St309, 36 y St2.

En conclusión, las cepas Cva53, Cr1, y En 8 presentarían resistencia otorgada por su *locus* CRISPR 3 aunque de manera menos eficiente que las cepas St307, St309, 36 y St2. Las cepas St307 y St309 serían las más resistentes ya que poseen dos de los espaciadores óptimos, mientras el resto de las cepas presentan solamente uno.

## 4. Resultados y discusión

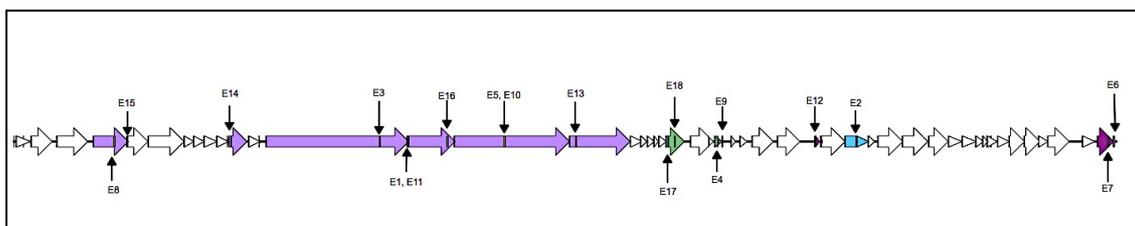


Figura 21. Representación esquemática de los espaciadores del locus CRISPR 3 en el genoma core del bacteriófago de *S.thermophilus* p107. Los espaciadores (E) están numerados del 1 al 18. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 16. Cuadro de referencias de la figura 21. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
CTCATATTCGTTAGTTGCTTTTGCATAAA	1	T, St175, Pot, Lf1, Hn, En24, En5, F81, Diag9, Diag12, 11L, Diag2	Rezagada	13164	Si	Si	No
TGATGGACGAGACGGTATTCCAGGAAAACC	2	St175, En24, En5, F81, 11L, Diag2	líder	28213	No	No	No
ATACTTACGATGGCGAAGATTACAACATATAG	3	St175, Lf1, En24, En5, F81, 11L, Diag2	líder	12247	Si	Si	No
TGATGAAAACAACATTACAGACTGGTGA	4	Fcia3, 3, Stsac2, 11CH, G72, Fcia2	líder	23583	No	Si	No
AGTCCAAAGGTATTTCCCTTAACCAAAGC	5	Fcia3, 3, Stsac2, 11CH, G72, Fcia2, 36, En6, St62	Rezagada	16426	Si	Si	No
TTATATAAAATCGTGCTACGGGCGTTTTAT	6	Diag9	Rezagada	36839	Si	Si	No
CTGAACCTAAAACGACTGATCTAACATCG	7	36, En6, St62	Líder	36649	No	No	No
AAGCGATGAAGCAAATGCAAGCGGTCCGT	8	36, En6, St62	Líder	3351	No	No	Si
TTTTATTTTTTTTAGTTTTCGTGGTCTAT	9	36, En6, St62	Rezagada	23718	No	Si	Si
GGTCCAAAGGTATTTCCCTTAATTAAGT	10	Fcia3, 3, Stsac2, 11CH, G72, Fcia2, 36, En6, St62	Rezagada	16426	Si	Si	No
CTCATATTCGTTAGTTGCTTTTGCATAAA	11	T, St175, Pot, Lf1, Hn, En24, En5, F81, Diag9, Diag12, 11L, Diag2	Rezagada	13164	Si	Si	No
CTTCTAAAATTGAATAATTGGAAGTACAT	12	St2	Líder	26840	Si	Si	No
TTGGATTGTTTGATTCCACCTTAGCACACT	13	St2	Rezagada	18819	Si	No	No
TATTAAGGTCCACAAAGCCTTCAACTG	14	TA	Líder	7260	Si	No	No
GGGTGAAAACCTAATAATACCACATTGA	15	TA	Líder	3775	Si	No	No
AAAATCATCTAAATACATGTGTGTAACAAG	16	TA	Rezagada	14507	No	Si	No
ATTGACCTATTCAATGTATGGGTACAGTAA	17	TA	Líder	21912	No	Si	No
GGGTGCTGATATGTCTACCTCTGGTGGAGC	18	TA	Líder	22128	Si	No	No

En la secuencia genómica del fago p107 se localizaron 18 espaciadores distribuidos en las siguientes 21 cepas diferentes; T, 11L, Diag2, St175, Pot, Lf1, Hn, Fcia3, Stsac2, 11CH, G72, Fcia2, En24, En5, F81, Diag9, 36, En6, St62, Diag12, St2 y TA.

De todos los espaciadores localizados E6, E7 y E12 están en regiones que codifican para proteínas hipotéticas.

Los espaciadores E2, E13, E14, E15 y E18 están situados en regiones conservadas del genoma y a su vez cumplen con todas las condiciones para brindar resistencia a las cepas que los contengan. Por lo tanto las cepas 11L, Diag2, St175, En24, En5, F81, St2 y TA presentarían una mayor resistencia ante el fago p107. Aunque la cepa TA sería la más resistente ya que contiene tres de los mejores espaciadores.

## 4. Resultados y discusión

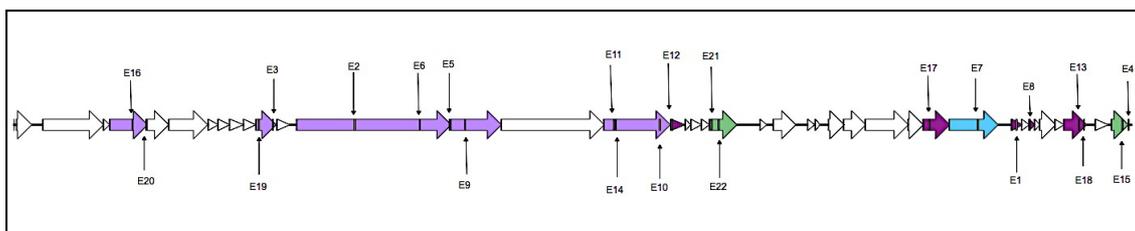


Figura 22. Representación esquemática de los espaciadores del locus CRISPR 3 en el genoma core del bacteriófago de *S. thermophilus* p128. Los espaciadores (E) están numerados del 1 al 22. En lila se representan las proteínas estructurales, en verde las proteasas, en celeste proteínas del módulo de procesamiento de ADN y en púrpura las proteínas hipotéticas.

Tabla 17. Cuadro de referencias de la figura 22. Se resaltan en color aquellos espaciadores capaces de brindar resistencia, en amarillo se marcan aquellos que se ubican en regiones poco variables y en lila aquellos que se encuentran en regiones variables

Espaciador	N° de espaciador	Cepas	Hebra	Posición	Mutación alternativa	Mutación secuencia semilla	Mutación PAM
TCTAAATCCAACAAGTCACAAAATCCTTT	1	Cva53, Cr1	Rezagada	31027	Si	No	No
TGTTTGGGAAACCGCAGTAGCCATGATTAA	2	Diag9, En24, Hn, Lf1, Pot, En5, F81, Lf2, Lf3, 11L	Líder	10536	No	No	No
TAACGGTGAGCCAATCAAAAAATACTTCAC	3	Diag9	Líder	8000	Si	No	Si
TTATATAAAATCGTGCTACGGCGTTTTAT	4	Diag9	Rezagada	34496	Si	Si	Si
CTCATATTCGTTAGTTGCTTTTGTCAAAA	5	Diag9, En24, Hn, Lf1, Pot, St175, T, 19, Diag5, EL, En5, En10, F81, Lf2, Lf3, 11L, TH4, Diag12, Diag2	Rezagada	13468	No	No	No
ATACTTACGATGGCGAAGATTACAACATATAG	6	En24, Lf1, St175, En5, F81, Lf2, Lf3, 11L, Diag2	Líder	12548	Si	Si	No
TGAATCTTCTAACTTTAACTCAGTTGTTAC	7	En24, Lf1, St175, St176, St179, En5, F81, Lf2, Lf3, Lf507, 11L, Lf502, Diag2	Líder	29826	Si	No	No
TATGCAAGTAAAGGAATATGCTTTATATAA	8	Hn, T, Diag5, En10, TH4	Líder	31434	No	No	No
AAACAAAGCCGAAGCCTTGGAAGTAGCGA	9	Lf1, Lf2, Lf3	Líder	13949	No	No	No
TATCTCTCTTTAACGACAGATGGTTTTTT	10	St307, St309	Rezagada	19982	Si	No	Si
GTACACGCTCACTTAATGGGTCGGGTGGT	11	St307, St309	Líder	18562	Si	No	No
TTCTAAGTCAAAAAGAAATATGATGCTAGTGG	12	St307, St309	Líder	20360	Si	No	No
CGGACAGCGATAAATACACTCTATACAGAGA	13	St307, St309	Líder	32961	No	Si	Si
AATCAGCAAATAAACAAGTACAACTATCCCA	14	En8	Líder	18619	No	No	No
CTGAACCTAAAACACTGACGTATCTAACATCG	15	36, En6, St62	Líder	34306	Si	No	Si
AAGCGATGAAGCAAATGCAAGGCGGTCCGT	16	36, En6, St62	Líder	3653	No	No	Si
TACATTCTTTGTCATTGACGTGGACATGCA	17	36, En6, St62	Líder	28334	No	No	No
TGATCGGATTTATTTAGAGTATCGCAC	18	St2	Líder	33116	Si	No	No
TATTAAGGTCCACAAAGCCTTTCAACTG	19	TA	Líder	7563	Si	No	No
GGGTGAAAACCTAATAATACCATTGA	20	TA	Líder	4077	Si	No	No
ATTGACCTATTCATGTATGGGTCACGTAA	21	TA	Líder	21586	No	Si	No
GGGTGCTGATATGCTACCTCTGGTGGAGC	22	TA	Líder	21803	Si	Si	No

En la representación genómica del fago p128 se ubicaron 23 espaciadores para CRISPR 3, distribuidos en las siguientes 32 cepas; Cva53, Diag9, En 24, Hn, Lf1, Pot, St175, St176, St179, T, 19, Cr1, Diag5, EL, En5, En10, 81, Lf2, Lf3, Lf507, St307, St309, En8, 11L, Diag2, TH4, 36, En6, St62, Diag12, Lf502, St2, Diag2 y TA.

De los 22 espaciadores encontrados E1, E8, E12, E13, E17 y E18 que se sitúan en regiones que codifican para proteínas hipotéticas.

Los espaciadores E2, E5, E7, E9, E11, E14, E17, E19 y E20 se ubican en regiones conservadas y también cumplen con los requisitos para brindar resistencia a las cepas que los contengan. Se trata de las cepas Diag9, En24, Hn, Lf1, Pot, En5, F81, Lf2, Lf3, 11L, Diag2, Diag9, En24, St175, T, 19, Diag5, EL, En10, TH4, Diag12, En24, St176, St179, Lf507, Lf502, St307, St309, En8, 36, En6, St62 y TA.

Las cepas Lf2, Lf3 y En 24 contienen tres de los mejores espaciadores, por lo tanto presentarían mejor resistencia ante el fago p128.

En la tabla 18 se compara la resistencia obtenida por métodos experimentales y bioinformáticos de las cepas ante cada fago, la información del *locus* en que está contenido el espaciador y si se encuentra ubicado en una región conservada o variable en el genoma del fago. Los espaciadores que se tomaron en cuenta, fueron aquellos que no presentaron mutaciones en la secuencia semilla ni en el motivo PAM.

#### 4. Resultados y discusión

Tabla 18. Comparativo entre los datos bioinformáticos y experimentales. Se indica en que *locus* se encuentra el espaciador adquirido (CRISPR 1 y/o CRISPR 3), (+) y (-) si el mismo se encuentra en una región conservada o variable respectivamente y en lila se señalan las cepas que muestran discrepancias entre los dos métodos.

Cepas	Datos bioinformáticos					Datos experimentales				
	Bacteriófagos					Bacteriófagos				
	p53	p73	p93	p107	p128	p53	p73	p93	p107	p128
TH4	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
CHE102	CRISPR 1 (+)	CRISPR 1 (+)	-	-	CRISPR 1 (+)	-	-	-	+	-
Diag2	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	-	+	-	-	-
En6	CRISPR 3 (-)	CRISPR 3 (-)	-	CRISPR 3 (-)	CRISPR 3 (+)	-	-	-	+	-
St307	-	CRISPR 3 (-)	CRISPR 3 (+)	-	CRISPR 3 (+)	+	-	-	-	-
St sac2	-	CRISPR 3 (+)	-	-	CRISPR 1 (+)	-	-	-	-	-
En8	CRISPR 1 (+)	CRISPR 1 (+)	CRISPR 3 (-)	-	CRISPR 1 (+)	-	-	-	+	-
					CRISPR 3 (+)					
Cva53	CRISPR 3 (-)	CRISPR 3 (-)	CRISPR 3 (-)	-	CRISPR 3 (-)	-	-	-	-	-
Fcia3	-	CRISPR 3 (+)	-	-	-	-	-	-	-	-
Lf 3	CRISPR 3 (+)	CRISPR 3 (+)	-	-	CRISPR 3 (+)	-	-	-	-	-
T	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
Diag 12	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	-	-	-	-	-
LF 507	-	-	-	-	CRISPR 3 (+)	-	-	-	-	-
St 2	-	CRISPR 3 (-)	CRISPR 3 (+)	CRISPR 3 (+)	CRISPR 3 (-)	-	-	-	-	-
19	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
11 CH	-	CRISPR 3 (+)	-	-	-	-	-	-	-	-
G 72	-	CRISPR 3 (+)	-	-	-	-	-	-	-	-
Cr1	CRISPR 1 (+)	CRISPR 3 (-)	CRISPR 3 (-)	CRISPR 1 (+)	CRISPR 1 (+)	+	-	-	+	-
	CRISPR 3 (-)				CRISPR 3 (-)					
Diag 9	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
St62	CRISPR 3 (-)	-	-	CRISPR 3 (-)	CRISPR 3 (+)	+	-	-	-	-
Pot 3	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
	-	CRISPR 3 (+)	-	-	-	-	-	-	+	-
En24	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	-	-	-	+	-
St 179	-	-	-	-	CRISPR 3 (+)	-	-	-	+	-
St43	CRISPR 1 (-)	-	-	-	CRISPR 1 (+)	-	-	-	+	-
Lf1	CRISPR 3 (+)	CRISPR 3 (+)	CRISPR 1 (+)	-	CRISPR 3 (+)	-	-	-	-	-
Lf2	CRISPR 3 (+)	CRISPR 3 (+)	-	-	CRISPR 3 (+)	-	-	-	-	-
Fcia2	-	CRISPR 3 (+)	-	-	-	-	-	-	-	-
St309	-	CRISPR 3 (-)	CRISPR 3 (+)	-	CRISPR 3 (+)	-	-	-	-	-
Hn	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	-	-	-	-	-
Lf502	-	-	-	-	CRISPR 3 (+)	-	-	-	-	-
TA	CRISPR 3 (+)	CRISPR 3 (+)	-	CRISPR 3 (+)	CRISPR 3 (+)	-	-	-	-	-
St176	-	-	-	-	CRISPR 3 (+)	-	-	-	-	-
En5	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	+	+	-	+	-
Diag 5	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
F 81	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	-	+	-	-	-
11L	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	-	+	-	-	-
EL	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
St175	CRISPR 3 (+)	-	-	CRISPR 3 (+)	CRISPR 3 (+)	-	-	-	-	-
En10	CRISPR 3 (+)	-	-	-	CRISPR 3 (+)	+	-	-	-	-
36	CRISPR 3 (-)	CRISPR 3 (-)	CRISPR 3 (+)	CRISPR 3 (-)	CRISPR 3 (+)	-	-	-	-	+
Tma	CRISPR 1 (-)	-	-	-	CRISPR 1 (+)	-	+	-	-	-

Analizando la tabla anterior se encuentran discrepancias para el bacteriófago p53 entre los resultados obtenidos experimentalmente y por análisis informáticos. El cuarto espaciador para p53 presenta tres mutaciones y según Deveau 2008 basta con que presente una mutación para inactivar el espaciador [30]. Este espaciador está presente en las cepas TH4, T, 19, Diag9, Pot, En5, Diag5 y En10. En las demás cepas que no presentan mutaciones, el espaciador se ubica en la hebra “antiparalela” afectando también la actividad del espaciador, de esta manera se explicaría porque son resistentes al bacteriófago experimentalmente (tabla 18) [30].

Para el caso del bacteriófago p107 el espaciador de la cepa Cr1 presenta mutaciones, por ende este se encontraría inactivo [30] y los demás pertenecen al *locus* CRISPR 3. una explicación podría ser una alta concentración de endolisinas fágicas en las suspensiones virales, que podrían estar ocasionando falsos positivos en los *spot* test. Se realizarán nuevos experimentos para confirmar que este no sea el caso.

La única inconsistencia real que se encontró son los espaciadores de la cepa 36 con respecto al bacteriófago p128.

Analizando ahora la actividad de los *loci*, en el caso de la cepa TH4 su *locus* activo parecería ser CRISPR 3, pero se sabe que en realidad CRISPR 1 está activo (Achigar, datos no publicados). Si bien se hallaron espaciadores en CRISPR 1 para todos los bacteriófagos, estos presentaban el motivo PAM o la región semilla mutada y por ende no podrían adjudicarle resistencia a la cepa. Para la cepa CHE102 no fue posible obtener el *locus* CRISPR 3 y CRISPR 1 está incompleto. CRISPR 1 podría ser el que se encuentra activo ya que ambos perfiles (experimental y bioinformático) de resistencia son similares, no se encontró mediante el análisis de los espaciadores resistencia ante el fago p93 y esto se podría ser explicado porque el *locus* no está completo.

Para la cepa En6 se encontraron espaciadores en CRISPR 3 capaces de hacerla resistente ante el fago p107 y como se detalló anteriormente no concuerda con lo observado experimentalmente, con lo cual puede ocurrir que CRISPR 1 sea el *locus* activo, pero al no contar con su secuencia para esta cepa, no lo podemos afirmar.

En el caso de la cepa En8 hay correlación entre ambos resultados y CRISPR 1 parecería ser el *locus* activo.

Para la cepa Lf3, parecería que CRISPR 3 fuera su *locus* activo, a pesar de no encontrar espaciadores adecuados contra los bacteriófagos p93 y p107.

Para la cepa Diag12, la secuencia para el *locus* CRISPR 1 no fue posible obtenerla y la secuencia de CRISPR 3 está incompleta. Según los datos obtenidos hasta el momento los perfiles de

resistencia van coincidiendo, y CRISPR 3 podría ser el que se encuentre activo, sin embargo, es necesario contar con la información completa para poder afirmarlo.

Para la cepa St2 los dos *locus* se encuentran incompletos. Se podría predecir que CRISPR 3 es su *locus* activo ya que se encontraron espaciadores capaces de conferir resistencia a los bacteriófagos para los cuales se comprobó experimentalmente que es resistente.

La cepa St62 presenta ambos *locus* incompletos. Aunque CRISPR 3 presenta espaciadores, los perfiles de resistencia no concuerdan, con lo cual CRISPR 1 podría ser el que se encontrara activo.

Para la cepa St43, el *locus* CRISPR 1 está incompleto y al *locus* CRISPR 3 no fue posible obtenerlo. Con lo cual, aunque CRISPR 1 parecería ser el que se encuentra activo, esto no se puede afirmar ya que es necesario complementar la información con los datos faltantes.

El *locus* CRISPR 3 está completo para la cepa Lf1 e incompleto para CRISPR 1. La cepa Lf2 también presenta CRISPR 3 completo, pero no se pudo obtener el *locus* CRISPR 1. Hay una gran concordancia entre los resultados obtenidos por ambos métodos, con lo cual CRISPR 3 puede ser el *locus* activo. Esto se puede reforzar ya que se ubican en el mismo *cluster* en CRISPR 3, diferenciándose solo en un espaciador y comparten el mismo perfil de resistencia.

Con respecto a la cepa TA, sus dos *locus* están incompletos, lo que explicaría por qué no se encontraron espaciadores en CRISPR 3 contra el fago p93, que es la única diferencia que se observa en las tablas. Por ende se podría establecer a CRISPR 3 como el *locus* activo.

Las cepas 11L, 19 y EL presentan el mismo *locus* CRISPR 1 mientras que las cepas 19 y EL tienen el mismo patrón de resistencia la cepa 11L es diferente (tabla 7). Respecto al *locus* CRISPR 3, las cepas 19, EL y 11L se encuentran en el mismo *cluster* (figura 12). Las cepas 19 y EL se diferencian solamente en un espaciador, con lo cual sería coherente que sean resistentes a los mismos bacteriófagos. Por otro lado cepa 11L se diferencia en una mayor cantidad de espaciadores, pudiendo presentar un patrón de resistencia diferente. Dado que el CRISPR 1 es idéntico el *locus* CRISPR 3 sería el que proporciona el patrón de resistencia distinto, siendo el más variable.

Las cepas En5 y F81 tienen ambos *locus* idénticos pero difieren en el patrón de resistencia (tabla 7). La cepa En 24 presenta el mismo *locus* CRISPR 3 que las cepas En5 y F81, y se ubica en el mismo *cluster* en el *locus* CRISPR 1, pero su secuencia no se pudo obtener en su totalidad. Mediante los resultados obtenidos por análisis bioinformático (tabla 18) se podría estimar al *locus* CRISPR 3 como activo, ya que ahí se encuentran los espaciadores capaces de brindarle resistencia según los criterios de Semenova [29] que hemos considerado y explicaremos más

adelante. Aunque si en efecto CRISPR 3 fuera el activo, las cepas tendrían que presentar el mismo patrón de resistencia, lo que no se observa mediante los datos experimentales (tabla 8). Lo que puede estar ocurriendo con la cepa En5 es que el *locus* no esté funcionando por algún motivo, por ejemplo por presentar mutaciones en alguna de las enzimas necesarias para su funcionamiento.

Las cepas En10 y Pot presentan un mismo *locus* CRISPR 1 y se ubican en el mismo *cluster* en CRISPR 3. Presentan resistencia ante todos los bacteriófagos excepto contra p53 (tabla 8) y los espaciadores adecuados se ubican en el *locus* CRISPR 3 (tabla 18) que en este caso es el más variable y activo.

Las cepas En7 y 36 tienen el mismo *locus* CRISPR 1. No podemos compararlas con el *locus* CRISPR 3 ya que En 7 no fue posible obtener su secuencia completa. Para la cepa 36 según el estudio de los espaciadores el *locus* CRISPR 3 sería el más variable y activo (tabla 18). Según datos experimentales el *locus* activo para la cepa En7 es CRISPR 1 [28]

Para las cepas St176, St179, Lf502 y Lf507 su *locus* CRISPR 3 es idéntico. Para las cepas St176 y Lf502 no fue posible obtener su *locus* CRISPR 1 con lo cual no se puede predecir que *locus* será el que se encuentre activo. Las cepas St179 y Lf507 se ubican en diferentes *clusters* en CRISPR 1. La cepa St179 según los ensayos experimentales es resistente a los bacteriófagos p53, p73, p93 y p128. Esta cepa tiene espaciadores en el *locus* CRISPR 1 contra todos estos bacteriófagos, excepto para p128, pero todos presentan mutaciones en sus motivos PAM, con lo cual, según el criterio que hemos adoptado, los espaciadores encontrados en CRISPR 1 no serían capaces de otorgarle resistencia a esta cepa (tablas 8-11).

Las cepas Diag2, 11L y St175 su *locus* CRISPR 3 es idéntico. Las cepas 11L y St175 se encuentran en el mismo *cluster* en el *locus* CRISPR 1. Sin embargo para la cepa Diag2 no fue posible obtener este *Locus*. Las cepas Diag2 y 11L presentan el mismo patrón de resistencia, y esto se observa tanto en los datos experimentales como bioinformáticos (tabla 18). La cepa St175 presenta el *locus* CRISPR 1 incompleto y ahí podría radicar la explicación a su patrón de resistencia diferente. Según los datos experimentales Diag2 y 11L son resistentes a cuatro bacteriófagos, mientras que St175 es resistente a cinco. Pero mediante el análisis bioinformático de estas tres cepas se encontró resistencia solo contra 3. Los espaciadores encontrados pertenecen al *Locus* CRISPR 3, por ende este podría ser el que se encontrase activo.

Las cepas TH4, Hn, Diag5, En10 y T presentan el mismo *locus* CRISPR 3. Por otro lado las cepas TH4, En10 y T se encuentran en el mismo *cluster* en CRISPR 1 y para las cepas Diag5 y Hn no

fue posible obtener la secuencia para este *locus*. Exceptuando Hn el resto de las cepas comparten el mismo perfil de resistencia y el *locus* CRISPR 3 es el que presentaría los espaciadores capaces de otorgársela (tabla 18). El *locus* CRISPR 1 es el que presenta mayor variabilidad. Para el caso de Hn su patrón de resistencia es diferente, y también se encontraron espaciadores en el *locus* CRISPR 3, pero al no tener secuenciado el CRISPR 1 no se puede inferir cual estará activo. En teoría tendría que ser este último, porque debería ser el más variable. En el caso de la cepa TH4 se comprobó experimentalmente que CRISPR 3 es su *locus* activo.

Las cepas Cva53 y Cr1 presentan igual *locus* CRISPR 3 y varían en su perfil de resistencia. Para la cepa Cva53 el *locus* CRISPR 3 presenta espaciadores capaces de otorgar resistencia, aunque éstos se ubican en regiones variables en el genoma de los bacteriófagos. Por otra parte, no es posible saber que ocurre con respecto a CRISPR 1 ya que no se obtuvo la secuencia. La cepa Cr1 tiene espaciadores tanto en CRISPR 1 como en CRISPR 3, pero los de CRISPR 1 se localizan en regiones poco variables en el genoma de los bacteriófagos, por lo tanto serían más estables. Por esta razón CRISPR 1 sería el *locus* activo, pero para afirmar esto sería necesario contar con su secuencia para la cepa Cva53.

Las cepas St307 y St309 tienen el mismo *locus* CRISPR 3 pero no se sabe lo que ocurre con respecto a CRISPR 1 ya que no fue posible obtenerlo para la cepa St307. Para ambas cepas se encontraron espaciadores contra bacteriófagos en CRISPR 3 (tabla 18) pero ya que difieren en el patrón de resistencia, en teoría el *locus* activo tendría que ser CRISPR 1 ya que debería ser el más variable.

Las cepas Fcia3, 3, Stsac2, 11CH, G72 y Fcia2 tienen el mismo *locus* CRISPR 3 y comparten el mismo *cluster* en CRISPR 1, excepto la cepa 3 para la que no fue posible obtenerlo. Ya que solamente la cepa 3 difiere en el patrón de resistencia, se podría suponer que CRISPR 1 es el que se encontraría activo, aunque se encuentran espaciadores en el *locus* CRISPR 3 capaces de brindar resistencia a las cepas (tabla 18).

Por último en la cepa Tma el *locus* CRISPR 1 se encuentra incompleto y a CRISPR 3 no fue posible obtenerlo. Con lo cual, aunque se encontraron espaciadores en CRISPR 1 y esos resultados coinciden con lo obtenido experimentalmente, al no contar con el *locus* CRISPR 3 no se puede decir con precisión cual se encontraría activo.

En la tabla 19 están representados aquellos espaciadores que se consideran mejores para ser utilizados para el mejoramiento de cepas de *S. thermophilus*, ya que están ubicados en regiones conservadas del genoma de los bacteriófagos, no tienen mutaciones en su región semilla ni en su secuencia PAM y aunque algunas presentan mutaciones puntuales (menos de 4)

#### 4. Resultados y discusión

se encuentran fuera de estas regiones, permitiéndole a un mismo espaciador reconocer bacteriófagos semejantes sin necesidad de aumentar el tamaño del *cassette* [29].

Tabla 19: Espaciadores candidatos para el mejoramiento de cepas *S.thermophilus*. (A) Para el *locus* CRISPR 1 y (B) para el *locus* CRISPR 3. En lila se resaltan los espaciadores que se encontraron dos veces en el análisis, en naranja los que se encontraron 4 veces y sin marcar aquellos que no repiten.

<b>Locus CRISPR 1</b>	
<b>A</b>	GATGTAACCCAAACCTTTTTGAGTTCTAC
	TGCTGAACGTATCTGTCCACTGTGTGGCCA
	TGGTTCTACACCTCAAACGGTGAGTGCATG
	CCAAATTTGCATTAACAAAACGCTCCTTC
	AGAAGCACCTCTTGC GTTGATAAAAAGTATT
	AATGCTGTAAAACCACGGGCTCGCTTAAGA
	ATACGCCATTAGGACTGAAAAGCATAGAA
	AGGCAGAAAAGAAGTATTTTGGTAAGTATG

<b>Locus CRISPR 3</b>	
<b>B</b>	CTCATATTCGTTAGTTGCTTTTGCATAAA
	AAACAAAGCCGAAGCCTTGTAAGTAGCGA
	TATTAAGGTCCACAAAGCCTTTCAACTG
	TATTAAGGTCCACAAAGCCTTTCAACTG
	GGGTGAAAACCCTAATAATACCACATTGA
	ATCAATTCATCTTGTATGTCTTCAATTGC
	TGTTTGGGAAACCGCAGTAGCCATGATTAA
	TGAATCTTCTAACTTTAACTCAGTTGTTAC
	GTACACGCTCACTTTAATGGGTCGGGTGGT
	AATCAGCAAATAACACTGACAACATCCCA
	TACATTCCTTGTCAATTGACGTGGACATGCA
	TATTAAGGTCCACAAAGCCTTTCAACTG
	AACGACTGCACAGCGGCAGTATGGGTGAAT
	GTATTTTTTACAAATGGCTAACATGAACTT
	TGATGGACGAGACGGTATTCCAGGAAAACC
	TTGGATTGTTTATTCCACCCTAGCACACT
	GGGTGCTGATATGTCTACCTCTGGTGGAGC

## **5. CONCLUSIONES Y PERSPECTIVAS**

A partir de las 50 cepas de *S. thermophilus* de la colección se lograron secuenciar los *loci* completos CRISPR 1 para 10 cepas y CRISPR 3 para 32 cepas y secuenciar parcialmente los *loci* CRISPR 1 para 25 cepas y CRISPR 3 para 6 cepas.

Se caracterizaron las cepas según el contenido de los espaciadores de los *locus* CRISPR 1 y CRISPR 3, agrupándolas en 11 *Clusters* para CRISPR 1 (figura 11) y en 6 *Clusters* para CRISPR 3 (figura 12).

Se ubicaron en las secuencias genómicas de los bacteriófagos uruguayos: p53, p93, p73, p128 y p107, los espaciadores de CRISPR 1 y CRISPR 3 para todas las cepas (figura 13-22). Tomando como criterio lo descrito por Semenova, se seleccionaron los potenciales espaciadores capaces de otorgar resistencia.

Mediante análisis comparativos de los datos bioinformáticos y experimentales de los patrones de resistencia (tabla 18) se predijo cual o cuales de los *locus* están activos para cada cepa de *S. thermophilus* estudiada.

Se seleccionaron aquellos espaciadores que se consideran mejores para ser utilizados en el mejoramiento de cepas de *S. thermophilus*. Los criterios considerados fueron los siguientes: espaciadores ubicados en regiones conservadas del genoma de los bacteriófagos, no poseen mutaciones en su región semilla ni en su secuencia PAM y aunque algunas presentan mutaciones puntuales (menos de 4) se encuentran fuera de estas regiones, permitiéndole a un mismo espaciador reconocer bacteriófagos semejantes sin necesidad de aumentar el tamaño del *cassette*. Al observarse inconsistencias entre los datos bioinformáticos y experimentales, el criterio elegido de Semenova 2011 no sería el más adecuado para nuestro estudio.

Como perspectiva de trabajo se deberían diseñar *primers* internos para aquellas cepas para las que no se lograron secuenciar los *locus* CRISPR de manera de completar adecuadamente la caracterización de las cepas de *S. thermophilus* de la colección estudiada.

## **6. BIBLIOGRAFÍA**

- [1] S. Özcelik, E. Kuley, and F. Özogul, "Formation of lactic, acetic, succinic, propionic, formic and butyric acid by lactic acid bacteria," *LWT - Food Sci. Technol.*, vol. 73, pp. 536–542, 2016.
- [2] S. Lahtinen, A. C. Ouwehand, S. Salminen, and A. von Wright, *Lactic Acid Bacteria: Microbiological and Functional Aspects, Fourth Edition*. Taylor & Francis, 2011.
- [3] F. Leroy and L. De Vuyst, "Lactic acid bacteria as functional starter cultures for the food fermentation industry," *Trends Food Sci. Technol.*, vol. 15, no. 2, pp. 67–78, 2004.
- [4] Y. Widyastuti and A. Febrisiantosa, "The role of lactic acid bacteria in milk fermentation," *Food Nutr. Sci.*, vol. 5, no. 4, p. 435, 2014.
- [5] Y. Widyastuti, Rohmatussolihat, and A. Febrisiantosa, "The Role of Lactic Acid Bacteria in Milk Fermentation," *Food Nutr. Sci.*, vol. 5, no. February, pp. 435–442, 2014.
- [6] R. Iyer, S. K. Tomar, T. Uma Maheswari, and R. Singh, "Streptococcus thermophilus strains: Multifunctional lactic acid bacteria," *Int. Dairy J.*, vol. 20, no. 3, pp. 133–141, 2010.
- [7] P. Hols *et al.*, "New insights in the molecular biology and physiology of Streptococcus thermophilus revealed by comparative genomics," *FEMS Microbiol. Rev.*, vol. 29, no. 3, pp. 435–463, 2005.
- [8] A. Quiberoni, S. Moineau, G. M. Rousseau, J. Reinheimer, and H.-W. Ackermann, "Streptococcus thermophilus bacteriophages," *Int. Dairy J.*, vol. 20, no. 10, pp. 657–664, 2010.
- [9] P. Horvath *et al.*, "Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus," *J. Bacteriol.*, vol. 190, no. 4, pp. 1401–1412, 2008.
- [10] R. W. Hendrix, "Bacteriophage genomics," *Curr. Opin. Microbiol.*, vol. 6, no. 5, pp. 506–511, 2003.
- [11] A. M. Q. King, *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*, vol. 9. Elsevier, 2011.
- [12] R. Calendar, *The bacteriophages*. Oxford University Press on Demand, 2006.
- [13] T. M. Michael, M. M. John, and P. Jack, "Brock microbiology of microorganism," *New Jersey. ISBN*, vol. 10, p. 130662712, 2002.
- [14] M. L. Capra *et al.*, "Extreme thermal resistance of phages isolated from dairy samples: Updating traditional phage detection methodologies," *Int. Dairy J.*, vol. 30, no. 2, pp.

- 59–63, 2013.
- [15] C. Le Marrec *et al.*, “Two groups of bacteriophages infecting *Streptococcus thermophilus* can be distinguished on the basis of mode of packaging and genetic determinants for major structural proteins.,” *Appl. Environ. Microbiol.*, vol. 63, no. 8, pp. 3246–3253, 1997.
- [16] M. B. Marcó, S. Moineau, and A. Quiberoni, “Bacteriophages and dairy fermentations,” *Bacteriophage*, vol. 2, no. 3, pp. 149–158, 2012.
- [17] R. Sorek, C. M. Lawrence, and B. Wiedenheft, “CRISPR-mediated adaptive immune systems in bacteria and archaea.,” *Annu. Rev. Biochem.*, vol. 82, pp. 237–66, 2013.
- [18] A. Kupczok and J. P. Bollback, “Probabilistic models for CRISPR spacer content evolution.,” *BMC Evol. Biol.*, vol. 13, p. 54, 2013.
- [19] K. S. Makarova *et al.*, “An updated evolutionary classification of CRISPR-Cas systems.,” *Nat. Rev. Microbiol.*, vol. 13, no. 11, pp. 722–736, 2015.
- [20] L. A. Marraffini and E. J. Sontheimer, “Invasive DNA, chopped and in the CRISPR,” *Structure*, vol. 17, no. 6, pp. 786–788, 2009.
- [21] N. Beloglazova *et al.*, “A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats,” *J. Biol. Chem.*, vol. 283, no. 29, pp. 20361–20371, 2008.
- [22] R. Barrangou, “The roles of CRISPR-Cas systems in adaptive immunity and beyond,” *Curr. Opin. Immunol.*, vol. 32, pp. 36–41, 2015.
- [23] D. Rath, L. Amlinger, A. Rath, and M. Lundgren, “The CRISPR-Cas immune system: Biology, mechanisms and applications,” *Biochimie*, vol. 117, pp. 119–128, 2015.
- [24] S. J. J. Brouns *et al.*, “Small CRISPR RNAs guide antiviral defense in prokaryotes,” *Science (80-. )*, vol. 321, no. 5891, pp. 960–964, 2008.
- [25] I. Yosef, M. G. Goren, and U. Qimron, “Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*,” *Nucleic Acids Res.*, p. gks216, 2012.
- [26] K. Chylinski, K. S. Makarova, E. Charpentier, and E. V. Koonin, “Classification and evolution of type II CRISPR-Cas systems,” *Nucleic Acids Res.*, vol. 42, no. 10, pp. 6091–6105, 2014.
- [27] A. Lopatina *et al.*, “CRISPR–Cas: Spacer Diversity Determines the Efficiency of Defense,” *Curr. Biol.*, vol. 26, no. 14, pp. R683–R685, 2016.

- [28] R. Achigar, "Mejoramiento para resistencia a bacteriófagos de cepas de *Streptococcus thermophilus* de uso en la industria láctea," 2014.
- [29] E. Semenova *et al.*, "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence," *Proc. Natl. Acad. Sci.*, vol. 108, no. 25, pp. 10098–10103, 2011.
- [30] H. Deveau *et al.*, "Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*," *J. Bacteriol.*, vol. 190, no. 4, pp. 1390–1400, 2008.